





Article

A New One-Parameter Distribution for Right Censored Bayesian and Non-Bayesian Distributional Validation under Various Estimation Methods

Walid Emam ¹, Yusra Tashkandy ¹, Hafida Goual ², Talhi Hamida ², Aiachi Hiba ², M. Masoom Ali ³, Haitham M. Yousof ^{4,*} and Mohamed Ibrahim ⁵

- ¹ Department of Statistics and Operations Research, Faculty of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia
- ² Department of Mathematics, Laboratory of Probability and Statistics LaPS, Badji Mokhtar Annaba University, Annaba 23000, Algeria
- ³ Department of Mathematical Sciences, Ball State University, Muncie, IN 47306, USA
- ⁴ Department of Statistics, Mathematics and Insurance, Faculty of Commerce, Benha University, Benha 13518, Egypt
- ⁵ Department of Applied, Mathematical and Actuarial Statistics, Faculty of Commerce, Damietta University, Damietta 34517, Egypt
- * Correspondence: haitham.yousof@fcom.bu.edu.eg

Abstract: We propose a new extension of the exponential distribution for right censored Bayesian and non-Bayesian distributional validation. The parameter of the new distribution is estimated using several conventional methods, including the Bayesian method. The likelihood estimates and the Bayesian estimates are compared using Pitman's closeness criteria. The Bayesian estimators are derived using three loss functions: the extended quadratic, the Linex, and the entropy functions. Through simulated experiments, all the estimating approaches offered have been assessed. The censored maximum likelihood method and the Bayesian approach are compared using the BB algorithm. The development of the Nikulin–Rao–Robson statistic for the new model in the uncensored situation is thoroughly discussed with the aid of two applications and a simulation exercise. For the novel model under the censored condition, two applications and the derivation of the Bagdonavičius and Nikulin statistic are also described.

Keywords: Bagdonavičius and Nikulin statistic; Bayesian estimation; BB method; censored applications; lomax model; Nikulin–Rao–Robson; Pitman's proximity

MSC: 62N01; 62N02; 62E10



Citation: Emam, W.; Tashkandy, Y.; Goual, H.; Hamida, T.; Hiba, A.; Ali, M.M.; Yousof, H.M.; Ibrahim, M. A New One-Parameter Distribution for Right Censored Bayesian and Non-Bayesian Distributional Validation under Various Estimation Methods. *Mathematics* **2023**, *11*, 897. <https://doi.org/10.3390/math11040897>

Academic Editor: Kai-Tai Fang

Received: 2 January 2023

Revised: 1 February 2023

Accepted: 3 February 2023

Published: 10 February 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

Exponential distribution is one of the often-encountered continuous probability distributions. It is frequently used to simulate the time between events. The mean and expected value of the exponential probability distribution, as well as its theoretical interpretation, will now be given. The exponential probability distribution, or Poisson point process, is the probability distribution of the interval between events in a process where events happen continuously and independently at a set average rate. In statistics and probability theory, this probability distribution is exploited. It is a unique instance of the gamma distribution. It is the continuous counterpart of the geometric distribution and possesses the key quality of being memoryless. In addition to the analysis of Poisson point processes, it is used in numerous other contexts. Exponential distribution is not the same as the class of exponential families of distributions, a major class of probability distributions that also includes the normal distribution, binomial distribution, gamma distribution, Poisson distribution, and many more probability distributions. One of the most often utilized continuous

distributions is exponential probability distribution. It helps in figuring out how long there will be between events. It is used in many disciplines, such as physics, dependability theory, queuing theory, and others. Finding the height of different molecules in a gas at a stable temperature and pressure in a uniform gravitational field, as well as computing the monthly and annual highest values of consistent rainfall and river outflow volumes, are all tasks that can be accomplished using the exponential probability distribution. In statistical inference and dependability, selecting a suitable core model for new data analysis is becoming increasingly crucial. The results could be significantly affected by even a little departure from the core model. The difficulty of this work is increased by censorship. The chi-square test type is the one used most frequently to determine the goodness-of-fit. Numerous modifications to chi-square tests have been proposed by various researchers.

One of the aims of this study is to present a goodness-of-fit test for our parametric model, which is often used in survival analysis, social sciences, engineering, and dependability, in complete data scenarios, and in the presence of right censoring. For the one parameter Poisson-exponential (OPPE) model, we present the explicit forms of the quadratic test statistics (Nikulin–Rao–Robson test, Bagdonavičius and Nikulin test statistic, see Bagdonavičius et al. [1], Bagdonavičius and Nikulin [2], Bagdonavičius and Nikulin [3], Nikulin [4], Nikulin [5], Nikulin [6], and Rao and Robson [7]). Then, we use actual data to apply both tests. This study demonstrates how the Bayesian technique may enhance both the mean time between failure of the OPPE distribution and the maximum likelihood estimate of the parameter. By using a suitable loss function, we will also illustrate the utility of this adjustment. The OPPE distribution is distinguished by its ease of mathematical and scientific handling in statistical and mathematical modeling. As is common for many probability distribution researchers, we do not approach this new distribution in the normal method in this work. For instance, we do not focus as much on the conventional study of the new distribution, not because it is unimportant, but rather because we are more concerned with the practical applications of the mathematical and statistical modeling, as well as a significant portion of the distribution's verification using censored data. To demonstrate the significance and flexibility of the new distribution and its wide range of applications in statistical and mathematical modeling, as well as the handling of controlled data, we omit several theoretical mathematical features, many algebraic derivations, and related theories. To be helpful to scholars in the field and assist them in providing additional similar and possibly more flexible distributions, we must briefly discuss the emergence of the new distribution in this context as well as how it was created and formed. The OPPE distribution's cumulative distribution function (CDF) can be written as

$$\Pr(X_\delta \leq x) = F_\delta(x) = \frac{1}{1 - \exp(-1)} \{1 - \exp[-\bar{\Delta}_\delta(x)]\},$$

where $x > 0$ and $\delta > 0$, $\bar{\Delta}_\delta(x) = 1 - \Delta_\delta(x)$ and

$$\Delta_\delta(x) = [\exp(-\delta x)]^{[1 - \exp(-\delta x)]}.$$

The corresponding probability density function (PDF) can then be expressed as

$$f_\delta(x) = \frac{\delta}{1 - \exp(-1)} \frac{\Delta_\delta(x)}{\exp\{[x + \bar{\Delta}_\delta(x)]\}} [\delta x + \exp(\delta x) - 1],$$

for all $x > 0$ and $\delta > 0$. Many scholars have been interested in the exponential distribution and in providing new flexible extensions of it, and they have also been interested in the applications of these new extensions in various fields of science, such as engineering, insurance, medicine, reliability, actuarial science, and others. In this work, we propose to develop a modified chi-square type fit test for the OPPE model in the scenario when the parameter is unknown, and the data are complete. This test is based on the Nikulin–Rao–Robson (NKRR) statistic, which Nikulin [3], Nikulin [4], Nikulin [5], Nikulin [6], and

Rao and Robson [7] separately proposed. This Y_n^2 statistic is a logical refinement of the conventional chi-square test based on the maximum likelihood estimator on the original data. The original data were used in full for this statistic, which is a straightforward enough statistic to estimate the null hypothesis's parameter. Second, this statistic entirely employed the Fisher data, which was to assess the information relative to a parameter contained in a model selection, based on the MLE. This provides a more unbiased evaluation of the chosen model of fit. Next, in the scenario when the parameters are unknown and the data are right-censored, we create a new goodness-of-fit test for this model. This modified chi-square Y_n^2 test adjusts the NKRR statistic to account for both censoring and the unknown parameter. This study is a very significant numerical simulation study that we are conducting to demonstrate the invariance of the test statistic distribution on the one hand, and to test the null hypothesis H_0 that a sample originates from an OPPE model on the other. We compute Y_n^2 the NKRR statistic as well as the Bagdonavičius and Nikulin statistics of various simulated samples, respectively. As a result, we can state that the suggested tests can, respectively, fit complete and censored data from the OPPE distribution. We conclude our contribution in this work with applications using real data, with the aim of demonstrating the applicability of the OPPE model across many scientific disciplines.

It is worth noting that the NKRR basic test, which is supported by complete data, is the most popular test in the last ten years. This is because it fits the complete truth data, and this is the case in most practical and applied cases in various fields. What distinguishes this statistical test also is the availability of ready-made statistical packages on the R program, whether for simulation or applications on actual data. However, the practical and experimental reality in many fields (such as the medical, chemical, engineering, etc.) necessitates that researchers deal with practical experiments that produce controlled censored data. This type of data, of course, needs certain tests dedicated to statistical dealing with it in the problem of distributional validation. The NKRR basic test is not the optimal choice in these cases. Hence, and based on this dilemma, was the primary and most important motive that prompted many researchers to think about introducing a new statistical test that fits the censored data. This new NKRR test, of course, is a modified test from the NKRR original test. According to the nature of the procedures of the two tests, both tests are not suitable for working with a type of data. One weakness of the original test, for instance, is that it is only applicable to entire data and cannot be used to deal with censored data. Therefore, this test may be a strong contender for doing statistical hypothesis tests if we are working with complete data. The modified test also has the drawback of only being appropriate for use with censored data; it is not designed for complete data and can only be used for censored data. Because it is designed for this type of data, this test will surely be a strong contender for statistical hypothesis tests if we are working with censored data.

2. Construction of NKRR Statistic for the OPPE Model

The NKRR statistic is a well-known variant of the traditional chi-squared tests in the situation of complete data (for more information, see Nikulin [3], Nikulin [4], Nikulin [5], Nikulin [6], and Rao and Robson [7]). The most common test to check whether a mathematical model is suitable for the data from observations is the chi-square statistic of Pearson. When the model's parameters are unknown or data are censored, these tests, however, cannot be used. Natural adaptations of the Pearson statistic for the entire set of data were reported by Nikulin [3], Nikulin [4], Nikulin [5], Nikulin [6], and Rao and Robson [7] and are known as NKRRs. The chi-square distribution is used in this statistical test, which is a logical extension of the Pearson statistic. When the censoring is included in addition to the unknown parameter, the classical test is insufficient to support the null hypothesis. Bagdonavičius et al. [1], Bagdonavičius and Nikulin [2], Bagdonavičius and Nikulin [3], Nikulin [4], Nikulin [5], Nikulin [6], and Rao and Robson [7] suggested changing the NKRR statistic to take into account random right censoring. For the OPPE model in the current study, we recommend creating a modified chi-square test. The following NKRR statistic Y^2

was developed by Nikulin [3], Nikulin [4], Nikulin [5], Nikulin [6], and Rao and Robson [7] to test the hypothesis.

$$H_0 : \Pr\{X_\delta \leq x\} = F_\delta(x) \Big|_{x \in R},$$

where in our case $F_\delta(x) = F_\delta(x)|_{x>0}$, x_1, x_2, \dots, x_n , is a sample belong to a parametric family $F_\delta(x)$, and where

$$Y^2(\hat{\delta}_n) = Z_n^2(\hat{\delta}_n) + n^{-1} L^T(\hat{\delta}_n) (I(\hat{\delta}_n) - J(\hat{\delta}_n))^{-1} L(\hat{\delta}_n),$$

where

$$Z_n^2(\hat{\delta}_n) = \left(\frac{\varsigma_1 - np_1(\hat{\delta}_n)}{\sqrt{np_1(\hat{\delta}_n)}}, \frac{\varsigma_2 - np_2(\hat{\delta}_n)}{\sqrt{np_2(\hat{\delta}_n)}}, \dots, \frac{\varsigma_n - np_b(\hat{\delta}_n)}{\sqrt{np_b(\hat{\delta}_n)}} \right)^T$$

and $J(\hat{\delta}_n)$ is the information (Inf-Mx) for the grouped data

$$J(\hat{\delta}_n) = \mathcal{B}(\hat{\delta}_n)^T \mathcal{B}(\hat{\delta}_n),$$

with

$$\mathcal{B}(\hat{\delta}_n) = \left[\frac{1}{\sqrt{p_i}} \frac{\partial p_i(\hat{\delta}_n)}{\partial \delta} \right]_{r \times s} \Big|_{(i=1,2,\dots,b \text{ and } k=1,2,\dots,s)},$$

and then

$$L(\delta) = (L_1(\delta), \dots, L_s(\delta))^T \text{ with } L_k(\delta) = \sum_{i=1}^r \frac{\varsigma_i}{p_i} \frac{\partial}{\partial(\delta)} p_i(\delta),$$

where $\hat{\delta}$ is the estimated Fisher Inf-Mx and $I(\hat{\delta})$ stands for the maximum likelihood estimator of the parameter vector. The Y^2 statistic has $(b-1)$ degrees of freedom and follows the χ_{b-1}^2 distribution. Consider a set of observations x_1, x_2, \dots, x_n that are collected in I_1, I_2, \dots, I_b (these b subintervals are mutually disjoint:

$$I_j = (a_{j,b}(x) - 1; x_{j,b}(x))$$

The intervals I_j 's limits for $a_{j,b}(x)$ are determined as follows

$$p_j(\delta) = \int_{a_{j,b}(x)-1}^{a_{j,b}(x)} f_\delta(x) dx \Big|_{(j=1,2,\dots,b)},$$

and

$$a_{j,b}(x) = F^{-1} \left(\frac{j}{b} \right) \Big|_{(j=1,\dots,b-1)}.$$

The vector of frequencies is generated by grouping the data into I_j intervals, $\varsigma_j = (\varsigma_1, \varsigma_2, \dots, \varsigma_b)^T$, where

$$\varsigma_j = \sum_{i=1}^n 1_{\{x_i \in I_j\}} \Big|_{(j=1,\dots,b)}.$$

In order to determine if the utilized data are distributed in accordance with the OPPE model, in the situation of an unknown parameter δ , we design an NKRR test statistic in this study as a modified goodness-of-fit test. We use the estimated Fisher Inf-Mx to provide all the components of the Y^2 statistic of our model after computing the maximum likelihood estimator $\hat{\delta}$ of the unknown OPPE distribution's parameter on the data set. For more applications under other different data sets, see Ibrahim et al. [8] and Yadav et al. [9].

3. Estimation and Inference

This section, via two subsections, discusses Bayesian and non-Bayesian estimating methods. Six non-Bayesian estimation methods are considered, including the maximum

likelihood estimation (MLE), the Cramér–von Mises estimation (CVME), the ordinary least square estimation (OLSQE), the weighted-least square estimation (WLSQE), moment method, and the Kolmogorov estimation (KE).

3.1. Classical Estimation Methods

3.1.1. Maximum Likelihood Method

Using some observable data, maximum likelihood estimation (MLE), a statistical approach, can estimate the unknown parameter of a probability distribution. Consider the n -sample (x_1, x_2, \dots, x_n) and a fixed constant m , we assume that the m -sample (x_1, x_2, \dots, x_m) is generated from the OPPE distribution. The likelihood function of this sample is

$$L_\delta(x) = \prod_{i=1}^m f_\delta(x_i) [1 - F_\delta(x_m)]^{n-m}$$

where $N = \frac{n!}{(n-m)!}$ in both (1) and (2) we have

$$L_\delta(x) = N\delta^m \delta^m (1 - \exp(-1))^{-n} \mathcal{C}_\delta(x_m)^{n-m} \prod_{i=1}^m \mathcal{A}_\delta(x_i) \mathcal{B}_\delta(x_i)$$

in which

$$\mathcal{A}_\delta(x_i) = \Delta_\delta(x_i) [\delta x_i - 1 + \exp(\delta x_i)],$$

$$\mathcal{B}_\delta(x_i) = \exp(-\delta x_i - (1 - \Delta_\delta(x_i))),$$

and

$$\mathcal{C}_\delta(x_m) = 1 - \exp(-(1 - \Delta_\delta(x_m))).$$

To obtain the maximum likelihood estimate (MLE) of δ , we have the log-likelihood ($\ell(\delta)$) function

$$\ell(\delta) = \log \left[\prod_{i=1}^n f_\delta(x_{i,n}) \right],$$

$$l_\delta(x) = \ln N + m \ln \delta - n \ln(1 - \exp(-1)) + (n - m) \ln \mathcal{C}_\delta(x_m) + \sum_{i=1}^m \ln \mathcal{A}_\delta(x_i) + \sum_{i=1}^m \ln \mathcal{B}_\delta(x_i).$$

The maximum likelihood estimator δ_{MLE} of the parameter δ is the solution of the following non-linear equation

$$\frac{\partial l_\delta(x)}{\partial \delta} = \frac{m}{\delta} + (n - m) \frac{1}{\mathcal{C}_\delta(x_m)} \frac{\partial \mathcal{C}_\delta(x_m)}{\partial \delta} + \sum_{i=1}^m \frac{1}{\mathcal{A}_\delta(x_i)} \frac{\partial \mathcal{A}_\delta(x_i)}{\partial \delta} + \sum_{i=1}^m \frac{1}{\mathcal{B}_\delta(x_i)} \frac{\partial \mathcal{B}_\delta(x_i)}{\partial \delta} = 0$$

where

$$\frac{\partial}{\partial \delta} \Delta_\delta(x_i) = -[x_i - x_i \exp(-\delta x_i) (1 - \delta x_i)] \exp(-\delta x_i (1 - \exp(-\delta x_i))),$$

$$\frac{\partial}{\partial \delta} \mathcal{A}_\delta(x_i) = \left(\frac{\partial}{\partial \delta} \Delta_\delta(x_i) \right) (\delta x_i - 1 + \exp(\delta x_i)) + \Delta_\delta(x_i) (x_i + x_i \exp(\delta x_i)),$$

$$\frac{\partial}{\partial \delta} \mathcal{B}_\delta(x_i) = \left[\left(\frac{\partial}{\partial \delta} \Delta_\delta(x_i) \right) - x_i \right] \exp(-\delta x_i - (1 - \Delta_\delta(x_i))),$$

and

$$\frac{\partial}{\partial \delta} \mathcal{C}_\delta(x_m) = - \left(\frac{\partial}{\partial \delta} \Delta_\delta(x_m) \right) \exp(-(1 - \Delta_\delta(x_m))).$$

3.1.2. The CVME Method

The CVME of the parameter δ is obtained via minimizing the following expression with respect to δ , where

$$\text{CVME}_{(\delta)} = \frac{1}{12}n^{-1} + \sum_{i=1}^n \left[F_{\delta}(x_{i,n}) - l_{(i,n)}^{[1]} \right]^2 |_{(x_{i,n} \in N_{(0)})},$$

and where $l_{(i,n)}^{[1]} = \frac{1}{2n}(2i-1)$, and where

$$\text{CVME}_{(\delta)} = \sum_{i=1}^n \left[\frac{1 - \exp[-\bar{\Delta}_{\delta}(x_{i,n})]}{1 - \exp(-1)} - l_{(i,n)}^{[1]} \right]^2.$$

Then, CVME of the parameter δ is obtained by solving the following non-linear equation

$$0 = \sum_{i=1}^n \left(\frac{1 - \exp[-\bar{\Delta}_{\delta}(x_{i,n})]}{1 - \exp(-1)} - l_{(i,n)}^{[1]} \right) \zeta_{(\delta)}(x_{i,n}, \delta),$$

where $\zeta_{(\delta)}(x_{i,n}, \delta) = \partial F_{\delta}(x_{i,n}) / \partial \delta$ is the first derivatives of the CDF of OPPE distribution with respect to δ .

3.1.3. The OLSQ Method

Let $F_{\delta}(x_{i,n})$ denote the CDF of the OPPE model and let $x_1 < x_2 < \dots < x_n$ be the n ordered RS. The OLSQE ($\mathcal{O}_{(\delta)}$) is obtained upon minimizing

$$\mathcal{OLS}_{(\delta)} = \sum_{i=1}^n \left[F_{\delta}(x_{i,n}) - l_{(i,n)}^{[2]} \right]^2,$$

where $l_{(i,n)}^{[2]} = \frac{i}{n+1}$. Then, we have

$$\mathcal{OLS}_{(\delta)} = \sum_{i=1}^n \left[\frac{1 - \exp[-\bar{\Delta}_{\delta}(x_{i,n})]}{1 - \exp(-1)} - l_{(i,n)}^{[2]} \right]^2.$$

The LSE is obtained via solving the following non-linear equation

$$0 = \sum_{i=1}^n \left[\frac{1 - \exp[-\bar{\Delta}_{\delta}(x_{i,n})]}{1 - \exp(-1)} - l_{(i,n)}^{[2]} \right] \zeta_{(\delta)}(x_{i,n}, \delta),$$

where $\zeta_{(\delta)}(x_{i,n}, \delta)$ is defined above.

3.1.4. The WLSQE Method

The WLSQE is obtained by minimizing the function $W_{(\delta)}$ with respect to δ

$$WLS_{(\delta)} = \sum_{i=1}^n d_{(i,n)}^{[3]} \left[F_{\delta}(x_{i,n}) - l_{(i,n)}^{[2]} \right]^2,$$

where $l_{(i,n)}^{[3]} = [(1+n)^2(2+n)]/[i(1+n-i)]$. The WLSQEs are obtained by solving

$$0 = \sum_{i=1}^n l_{(i,n)}^{[3]} \left[\frac{1 - \exp[-\bar{\Delta}_{\delta}(x_{i,n})]}{1 - \exp(-1)} - l_{(i,n)}^{[2]} \right] \zeta_{(\delta)}(x_{i,n}, \delta),$$

where $\zeta_{(\delta)}(x_{i,n}, \delta)$ is defined above.

3.1.5. Method of Moments

The method of moments is a technique used in statistics to estimate population parameters. To derive higher moments like skewness and kurtosis, the same method is applied. It begins by describing the population moments as functions of the important parameters (i.e., the anticipated powers of the random variable under discussion). The sample moments are then set to be equal to those. There are exactly as many of these equations as there are parameters that need to be estimated. Then, the equations are solved for the relevant parameters. Estimates of those parameters are used in the solutions. The moment estimation of the one-parameter of the OPPE distribution can be obtained by equating the first theoretical moment of (2) with the corresponding sample moments as follows

$$\mu'_1 = E(X_{i,n}) = \frac{1}{n} \sum_{i=1}^n x_{i,n}$$

where μ'_1 can be evaluated from (2).

3.1.6. KE Method

The Kolmogorov estimate (KE) $\hat{\delta}$ of δ is obtained by minimizing the function

$$KE = KE(\delta) = \max_{1 \leq i \leq n} \left\{ \frac{1}{n} i - F_{\delta}(x_{i,n}), F_{\delta}(x_{i,n}) - \frac{1}{n} (i-1) \right\}.$$

The KE $\hat{\delta}$ of δ is obtained by comparing $\left[\frac{1}{n} i - F_{\delta}(x_{i,n}) \right]_{1 \leq i \leq n}$ and $[F_{\delta}(x_{i,n}) - \frac{1}{n} (i-1)]_{1 \leq i \leq n}$ and selecting the max one. However, for $1 \leq i \leq n$, we are minimizing the whole function $K(\delta)$. For more detail about the KE method.

3.2. Simulations and Assessment

For comparing the classical methods, some MCMC simulation studies are performed. The results are presented in: Table 1 ($\delta = 0.5 \mid n = 50, 100, 200, 300$, and 500); Table 2 ($\delta = 0.9 \mid n = 50, 100, 200, 300$, and 500) and Table 3 ($\delta = 1.5 \mid n = 50, 100, 200, 300$, and 500). The numerical assessments are performed depending on the mean squared errors (MSEs). First, we generate $N = 1000$ samples of the OPPE model. Based on Tables 1–3, it is noted that the performance of all estimation methods improves when $n \rightarrow +\infty$. Despite the variety and abundance of the other classic methods, as demonstrated in Tables 1–3, the MLE approach is still the most efficient and reliable of the surviving classic methods. The MLE approach is generally mentioned as being advised for statistical modeling and applications. This evaluation is mostly based on a thorough simulation research, as displayed in Tables 1–3. Since the MSE for the MLE is the smallest for all $n = 50, 100, 200, 300$, and 500, this section uses simulation studies to evaluate rather than compare various estimating procedures. However, this does not exclude the use of simulation to compare various estimating approaches. Nevertheless, actual data are commonly used to assess different estimating approaches; for this reason, we will discuss a few examples specifically for this role. There are two more applications to compare the competing models.

Table 1. Simulation results for $\delta = 0.5$.

n	MLE	OLSQ	WLSQ	CVM	Moment	KE
50	0.00305	0.00334	0.00314	0.00355	0.00329	0.00366
100	0.00148	0.00163	0.00163	0.00161	0.00163	0.00179
200	0.00073	0.00081	0.00081	0.00080	0.00082	0.00088
300	0.00047	0.00054	0.00051	0.00054	0.00049	0.00055
500	0.00027	0.00031	0.00031	0.00031	0.00030	0.00032

Table 2. Simulation results for $\delta = 0.9$.

n	MLE	OLSQ	WLSQ	CVM	Moment	KE
50	0.00981	0.01073	0.01035	0.00939	0.01195	0.01187
100	0.00489	0.00533	0.00511	0.00556	0.00539	0.00616
200	0.00248	0.00286	0.00278	0.00282	0.00239	0.00285
300	0.00155	0.00177	0.00177	0.00178	0.00174	0.00180
500	0.00092	0.00098	0.000987	0.00099	0.00095	0.00108

Table 3. Simulation results for $\delta = 1.5$.

n	MLE	OLSQ	WLSQ	CVM	Moment	KE
50	0.02580	0.03189	0.02847	0.02888	0.03039	0.03448
100	0.01296	0.01591	0.01414	0.01509	0.01408	0.01582
200	0.00633	0.00749	0.00732	0.00744	0.00740	0.00785
300	0.00460	0.00466	0.00498	0.00480	0.00494	0.00504
500	0.00271	0.00299	0.00291	0.00298	0.00302	0.00301

3.3. Applications for Comparing Methods

Here, we will be very interested in comparing the different estimation methods, but this time through real data. The first data set called the failure time data or relief times (in minutes) of patients receiving an analgesic (see Gross and Clark [10]). The second data set is called the survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli, see Bjerkedal [11].

We discuss the skewness–kurtosis plot (or the Cullen and Frey plot) in these applications for examining initial fits of theoretical distributions such as normal, uniform, exponential, logistic, beta, lognormal, and Weibull. Plotting and bootstrapping are both applied for greater accuracy. The scattergram plots, the “nonparametric Kernel density estimation (NKDE)” method for examining the initial shape of the insurance claims density, the “Quantile-Quantile (Q-Q)” plot for examining the “normality” of the current data, the “total time in test (TTT)” plot for examining the initial shape of the empirical hazard rate function (HRF), and the “box plot” for identifying the extreme data were also presented.

Figure 1 gives the box plot for the failure time data (first row, the left panel), Q-Q plot for the failure time data (first row, the right panel), TTT plot for the failure time data (second row, the left panel), nonparametric Kernel density estimation plot for the failure time data (second row, the right panel), the Cullen and Frey plot for the failure time data (third row, the left panel), and scattergrams (third row, the right panel) for the failure time data.

Based on Figure 1 (first row), the relief data have only one extreme observation, based on Figure 1 (second row, the left panel), the HRF of the relief times is “monotonically increasing HRF”, based on Figure 1 (second row, the right panel), nonparametric Kernel density estimation is bimodal and right skewed with asymmetric shape, based on Figure 1 (third row, the left panel), the relief times data do not follow any of the theoretical distributions such as the normal, uniform, exponential, logistic, beta, lognormal, and Weibull. Figure 2 gives the box plot for the survival times (first row, the left panel), Q-Q plot for the survival times (first row, the right panel), TTT plot for the survival times (second row, the left panel), nonparametric Kernel density estimation plot (second row, the right panel), the Cullen and Frey plot (third row, the left panel), and scattergrams (third row, the right panel) for the survival times data. The survival data in Figure 2 (first row) has four extreme observations, the HRF of the survival times is “monotonically increasing HRF,” the nonparametric Kernel density estimation is bimodal and right skewed with an asymmetric shape, and the survival times data in Figure 2 (third row, left panel) do not follow any of the theoretical distributions, such as the normal or gamma distributions.

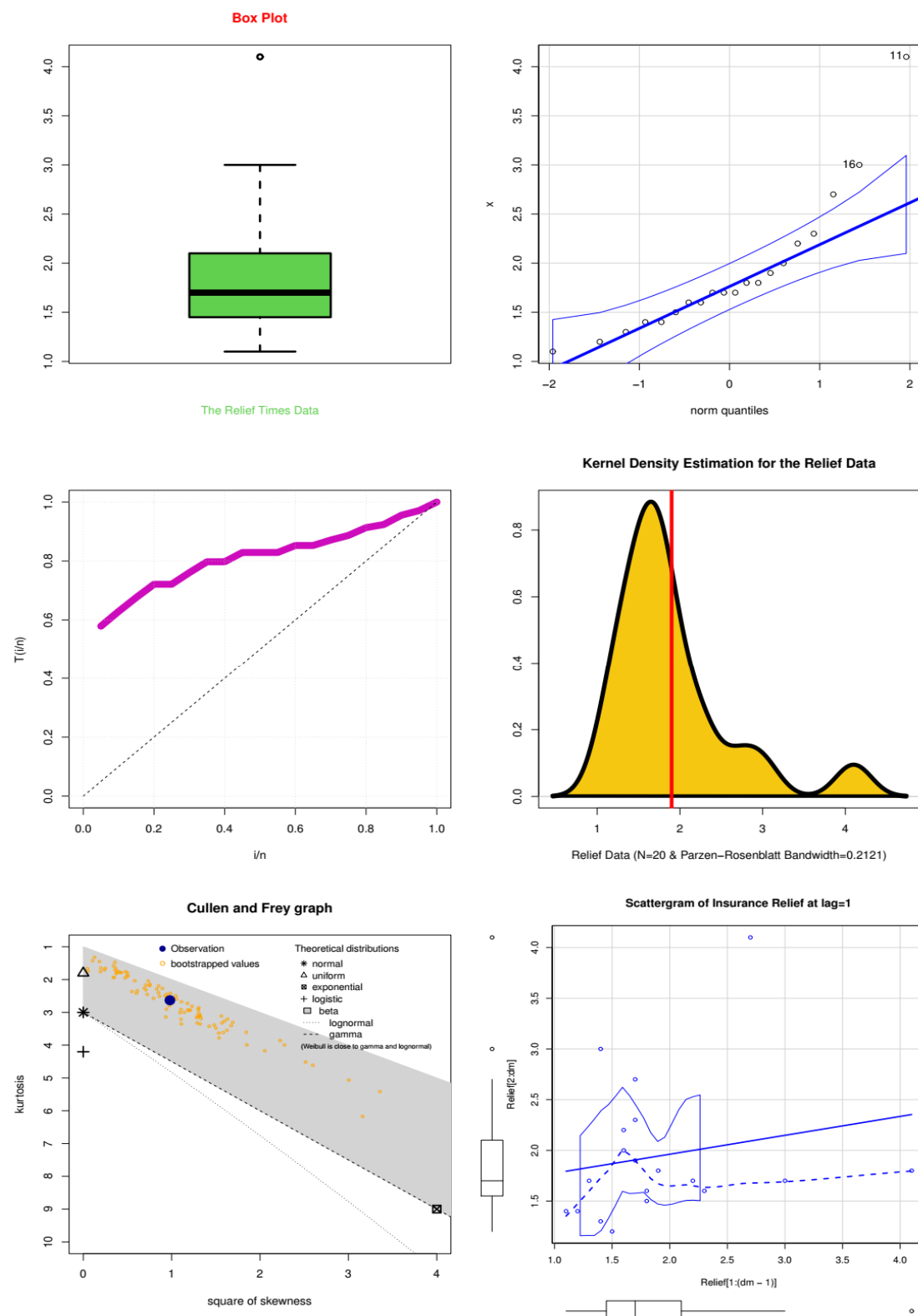


Figure 1. Box plot, Q-Q plot, TTT plot, nonparametric Kernel density estimation, Cullen and Frey, and scattergrams for the relief times.

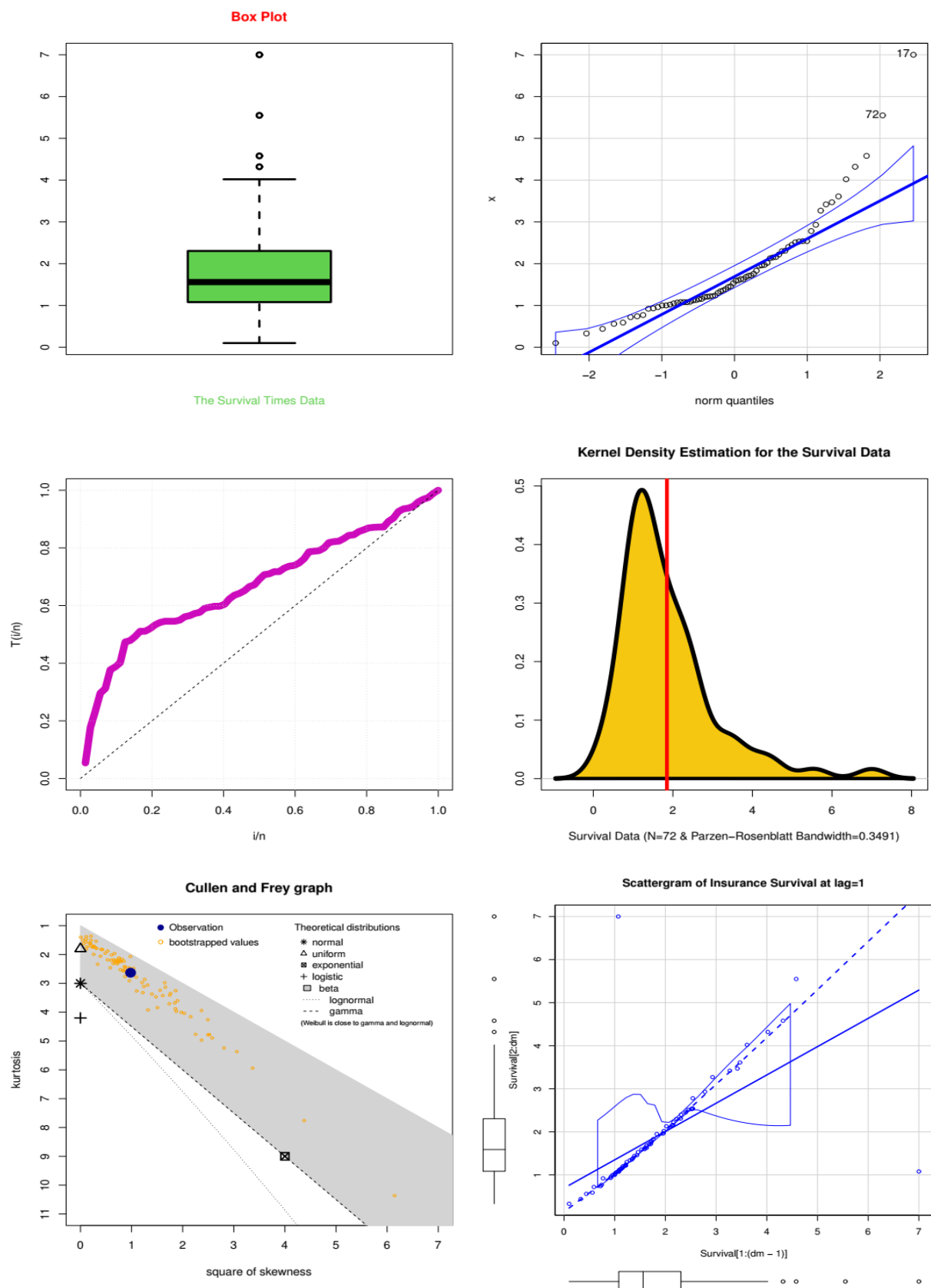


Figure 2. Box plot, Q-Q plot, TTT plot, nonparametric Kernel density estimation, Cullen and Frey, and scattergrams for the survival times.

A statistical test called the Anderson–Darling test (ADT) is used to determine if a sample of data is taken from a particular probability distribution. The test is based on the assumption that the distribution under the test has no parameters that need to be estimated, in which case the test and its set of critical values are distribution-free. The test is most frequently applied when a family of distributions is being examined, in which case it is necessary to estimate the family’s parameters and take this into consideration when modifying the test statistic or its critical values. It is one of the most effective statistical strategies for detecting most deviations from normal distribution when used to determine whether a normal distribution adequately describes a collection of data.

The Cramér–von Mises criterion (CVMC) in statistics is a criterion used to compare two empirical distributions or to determine how well a cumulative distribution function fits an empirical distribution function. In other techniques, such as minimum distance, the CVMC and ADT tests are used in comparing methods. Table 4 gives the application results (CVMC and ADT) for comparing methods under the relief data. Table 5 lists the application CVMC and ADT for comparing methods under the survival data. Based on Table 4, it is seen that the moment method is the best with CVMC = 0.08439 and ADT = 0.499304, then the ML method with W CVMC = 0.085876 and ADT = 0.508074. Based on Table 4, it is seen that the moment method is the best with CVMC = 0.070712 and ADT = 0.463826, then the ML method with CVMC = 0.071247 and ADT = 0.465521.

Table 4. *p*-values for comparing methods under the relief data.

Method	δ	CVMC	ADT
ML	0.503284	0.085876	0.508074
LS	0.470917	0.086923	0.514258
WLSQ	0.428757	0.088466	0.523383
CVM	0.470955	0.086921	0.514251
Moment	0.557676	0.08439	0.499304
KE	0.467910	0.087026	0.514869

Table 5. *p*-values for comparing methods under the survival data.

Method	δ	CVMC	ADT
ML	0.567783	0.071247	0.465521
LS	0.546370	0.071690	0.467139
WLSQ	0.551350	0.071581	0.466728
CVM	0.546576	0.071685	0.467122
Moment	0.599247	0.070712	0.463826
KE	0.527965	0.072124	0.468851

3.4. Bayesian Analysis under Different Loss Functions

In this section, we will review some statistical aspects related to Bayesian estimations. Three types of Bayesian loss functions have been used. Additionally, I made a lot of useful comparisons using the new distribution. Many Bayesian statistical algorithms can be utilized, as per Muñoz-Gil et al. [12], and within their particular context, they can aid researchers in mathematics and statistical modeling tasks, particularly when applying Bayes’ theory. Muñoz-Gil et al. [12] offered the proposed representation of four full algorithms with their codes and implementations for this specific purpose, allowing the reader to follow the new algorithms and present more algorithms based on them.

3.4.1. Prior and Posterior Distributions

As prior distributions, we assume the parameter δ has an informative distribution as a prior, which is a Gamma distribution:

$$\pi(\delta) = \frac{ab}{\Gamma(b)} \delta^{b-1} \exp(-a\delta), a, b > 0.$$

The posterior distribution of δ is

$$\pi(\delta|x) = K\delta^{m+b-1}[1 - \exp(-1)]^{-m}\exp(-a\delta)\prod_{i=1}^m P_{\delta}(x_i)\Delta_{\delta}(x_i)\mathcal{Q}_{\delta}(x_m)^{n-m},$$

where

$$K = \int_0^{+\infty} \delta^{m+b-1}[1 - \exp(-1)]^{-m}\exp(-a\delta)\prod_{i=1}^m P_{\delta}(x_i)\Delta_{\delta}(x_i)\mathcal{Q}_{\delta}(x_m)^{n-m}d\delta,$$

is the normalizing constant. In the following, we use the three loss functions, namely the generalized quadratic (GQ), the entropy, and the Linex loss functions to obtain the Bayesian estimators. In the following formulas, we consider γ , P , and r as integers.

3.4.2. Bayesian Estimators and Their Posterior Risk

The Bayesian estimator under the generalized quadratic loss function (GQ) is

$$\delta_{GQ} = \frac{I_0^{+\infty}(m+b-1+\gamma, x_m)}{I_0^{+\infty}(m+b-2+\gamma, x_m)},$$

where

$$I_0^{+\infty}(m+b-1+\gamma, x_m) = \int_0^{+\infty} \delta^{m+b-1+\gamma}[1 - \exp(-1)]^{-m}\exp(-a\delta)\prod_{i=1}^m \Delta_{\delta}(x_i)P_{\delta}(x_i)\mathcal{Q}_{\delta}(x_m)^{n-m}d\delta$$

and

$$I_0^{+\infty}(m+b-2+\gamma, x_m) = \int_0^{+\infty} \delta^{m+b-2+\gamma}[1 - \exp(-1)]^{-m}\exp(-a\delta)\prod_{i=1}^m \Delta_{\delta}(x_i)P_{\delta}(x_i)\mathcal{Q}_{\delta}(x_m)^{n-m}d\delta.$$

Under the entropy loss function, we obtain the following estimator

$$\delta_E = \left[K \int_0^{+\infty} \delta^{m+b-1-p}[1 - \exp(-1)]^{-m}\exp(-a\delta)\prod_{i=1}^m \Delta_{\delta}(x_i)P_{\delta}(x_i)\mathcal{Q}_{\delta}(x_m)^{n-m}d\delta \right]^{-\frac{1}{p}}.$$

The corresponding posterior risk is

$$PR(\delta_E) = PE_{\pi}(\ln(\delta) - \ln(\delta_E)),$$

finally, under the Linex loss function, the Bayesian estimator

$$\delta_L = \frac{-K}{r} \ln \left[\int_0^{+\infty} \delta^{m+b-1}[1 - \exp(-1)]^{-m}\exp[-\delta(a+r)]\prod_{i=1}^m P_{\delta}(x_i)\Delta_{\delta}(x_i)\mathcal{Q}_{\delta}(x_m)^{n-m}d\delta \right],$$

and the corresponding posterior risk is

$$PR(\delta_L) = r(\delta_{GQ} - \delta_L).$$

Since it is unlikely possible to obtain all these estimators analytically, we suggest the use of the MCMC procedures to evaluate them. In this work, several Bayesian algorithms were developed recently for the analysis of tracer-diffusion single-particle tracking data. Following Muñoz-Gil et al. [12], many Bayesian statistical algorithms can be used, which within their specific framework, can help researchers in mathematical and statistical modeling operations, especially when using Bayes' theory. For this specific purpose, Muñoz-Gil et al. [12] presented the proposed visualization of four complete algorithms

with their codes and applications, and the reader can track the new algorithms and other algorithms can be presented based on them.

3.4.3. Comparing the Likelihood Estimation and the Bayesian Estimation Using Pitman's Closeness Criterion

We are going to compare the performance of the proposed Bayesian estimators with the MLEs, for that purpose, we perform a MCMC simulation method with $\delta = 1.5$ and $a = 3$, $b = 2$. We generate $N = 1000$ type II censored samples following the OPPE model, we use different sample sizes $n = 30, 100, 200$ while $m = 10, 40, 160$, respectively, and we obtain the following results. Table 6 contains the values of the estimators produced by the function BB algorithm. Here, we observe that, especially as sample size n is raised, the estimated values are fairly close to the parameter's genuine values. Tables 7–9 present the Bayesian estimators and PR (in brackets) under the generalized quadratic loss function, the entropy loss function, and the Linex loss function, respectively. Table 10 lists the Bayesian estimators and PR (in brackets) for each of the three loss functions.

Table 6. The MLE of the parameter with quadratic error (in brackets).

N = 5000	n = 30	n = 100	n = 200
m →	10	40	160
$\delta \rightarrow$	1.4905 (0.0125)	1.46543 (0.0321)	1.5104 (0.0023)

Table 7. Bayesian estimators and PR (in brackets) under GQ loss function.

$\gamma \downarrow N = 5000, m \rightarrow$	n = 30, 10	n = 100, 40	n = 200, 160
−2	1.6490 (0.0076)	1.6825 (0.0053)	1.6432 (0.0023)
−1.5	1.7990 (0.0087)	1.0825 (0.0061)	1.2127 (0.0016)
−1	1.6182 (0.0005)	0.9739 (0.0001)	2.0018 (0.0001)
−0.5	1.4994 (0.0071)	1.4888 (0.0060)	1.5138 (0.0011)
0.5	1.9760 (0.0085)	1.7926 (0.0077)	1.3439 (0.0013)
1	1.7516 (0.0087)	1.2972 (0.0058)	1.7208 (0.0035)
1.5	1.6743 (0.0045)	1.5632 (0.0065)	1.3275 (0.0032)
2	1.4768 (0.1241)	1.4191 (0.1181)	1.7158 (0.0033)

Table 8. Bayesian estimators and PR (in brackets) under the entropy loss function.

N = 5000	n = 30	n = 100	n = 200
p ↓ m →	10	40	160
−2	1.6037 (0.0009)	0.7757 (0.3190)	1.8493 (0.0308)
−1.5	1.3990 (0.1644)	1.2144 (0.0019)	1.0942 (0.0080)
−1	1.0886 (0.0031)	0.7654 (0.1173)	1.7697 (0.0099)
−0.5	1.6001 (0.0065)	1.542 (0.00020)	1.5638 (0.0071)
0.5	1.483 (0.00413)	1.4899 (0.0023)	1.5121 (0.0012)
1	1.4579 (0.0997)	1.4354 (0.0944)	1.4571 (0.0014)
1.5	1.6981 (0.0038)	0.4830 (0.0733)	1.2148 (0.0009)
2	1.356 (0.16440)	1.0942 (0.0080)	1.2144 (0.0019)

In Table 7, the estimation under the generalized quadratic loss function, we remark that the value $\gamma = -1$ gives the best posterior risk. Additionally, we obtain the smallest suitable posterior risk when n is long. In the estimation under the entropy loss function, we obtain Table 8 where we can notice that the value $p = -0.5$ when $n = 200$ provides the best posterior risk. We can notice clearly that the value $r = -1.5$ provides the best PR. In conclusion, it is evident from a brief comparison of the three loss functions that the quadratic loss function yields the best outcomes; the results are further illustrated in Table 10. The best Bayesian estimators and maximum likelihood estimators should be

compared, as we suggest. We employ the Pitman closeness criterion for this purpose (more information can be found in Pitman [13], Fuller [14], and Jozani [15]).

Table 9. Bayesian estimators and PR (in brackets) under Linex loss function.

$r \downarrow$	$N = 5000$	$n = 30$	$n = 100$	$n = 200$
	$m \rightarrow$	10	40	160
−2		1.3815 (0.0182)	1.5434 (0.0183)	1.3815 (0.0001)
−1.5		1.2080 (0.0014)	1.2609 (0.0199)	1.4011 (0.0004)
−1		1.6234 (0.0108)	0.9900 (0.0103)	1.2819 (0.0065)
−0.5		1.1915 (0.0111)	1.3187 (0.0195)	1.4091 (0.0032)
0.5		1.3909 (0.0231)	1.4560 (0.0183)	1.3324 (0.0032)
1		1.3815 (0.0183)	1.3815 (0.0183)	1.3815 (0.0183)
1.5		0.5155 (0.0519)	0.5315 (0.0183)	1.5109 (0.0057)
2		1.4193 (0.0131)	0.9547 (0.1041)	1.4045 (0.0004)

Table 10. Bayesian estimators and PR (in brackets) under the three loss functions.

$N = 5000$	$n_1 = 10$	$n_2 = 50$	$n_3 = 200$
$m \rightarrow$	10	40	160
$GQ _{\gamma=1}$	1.4994 (0.00710)	1.4888 (0.0060)	1.5138 (0.0011)
$Entropy _{p=-0.5}$	1.4830 (0.00413)	1.4899 (0.0023)	1.5121 (0.0012)
$Linex _{r=-1.5}$	0.5155 (0.05190)	0.5315 (0.0183)	1.5109 (0.0057)

Definition 1. An estimator ϑ_1 of a parameter ϑ dominates another estimator ϑ_2 in the sense of Pitman's closeness criterion if, for all $\vartheta \in \Theta$,

$$P_{\vartheta}[|\vartheta_1 - \vartheta| < |\vartheta_2 - \vartheta|] > 0.5.$$

We show the Pitman probability' values in Table 11 so that we may contrast the Bayesian estimators with the MLE estimator when using the three loss functions when $\gamma = -1$, $p = -0.5$, and $r = -1.5$. Then, we notice that, according to this criterion, the Bayesian estimators of the parameter is better than the MLE. Additionally, the Linex loss function has the best values in comparison with the other two loss functions with the probability 0.734, $\delta_{(n=200, m=160)}$.

Table 11. Pitman comparison of the estimators.

$N = 5000$	$n = 10$	$n = 50$	$n = 200$
$m \rightarrow$	8	40	160
$GQ _{\gamma=1}$	0.678	0.654	0.674
$Entropy _{p=-0.5}$	0.534	0.587	0.632
$Linex _{r=-1.5}$	0.587	0.5789	0.734

4. Distributional Validation

We utilize the statistic type test based on a version of the NKRR statistic suggested by Bagdonavičius et al. [1], Bagdonavičius and Nikulin [2], Bagdonavičius and Nikulin [3], Nikulin [4], Nikulin [5], Nikulin [6], and Rao and Robson [7] to confirm the sufficiency of the OPPE model when the parameter is unknown, and the data are censored. The failure rate x_i follows an OPPE distribution, hence we modify this test for an OPPE model. Let's us think about the null hypothesis:

$$H_0 : F(x) \in F_0 = F_{0,\delta}(x)|_{x \in R},$$

The OPPE distribution's survival function (SrF) and cumulative hazard function are as follows:

$$S_{\delta}(x) = 1 - F_{\delta}(x) = 1 - \left[\frac{1}{1 - \exp(-1)} (1 - \exp\{-[1 - \Delta_{\delta}(x)]\}) \right],$$

where

$$\Delta_{\delta}(x) = [\exp(-\delta x)]^{[1 - \exp(-\delta x)]}.$$

and

$$V_{\delta}(x) = -\ln[S_{\delta}(x)],$$

For all j , we have a constant value of $e_{j,X} = E_k/k$ under this choice of intervals. Since the inverse hazard function of the OPPE distribution lacks an explicit form, intervals can be estimated iteratively. Let us divide a finite time period $[0, \tau]$ into $k > s$ smaller periods $I_j = (a_{j-1}(x), a_{j,b}(x)]$ and

$$0 = a_{0,b} < a_{1,b} \dots < a_{k-1,b} < a_{k,b} = +\infty.$$

If $x_{(i)}$ is the i th element in the ordered statistics $(x_{(1)}, \dots, x_{(n)})$, and if $[V_{\delta}(x)]^{-1}$ is the inverse of the cumulative hazard function $V_{\delta}(x)$, then the estimated value of $\hat{a}_{j,b}(x)$ where

$$E_k = E_k(x) = \sum_{i=1}^{i-1} V_{\delta}(x),$$

and $a_{j,b}(x)$ are random data functions, such as the k selected intervals have equal expected numbers of failures $e_{j,X}$. The test for hypothesis H_0 can be based on the statistic

$$Y_{n,r-1,\varepsilon}^2(\hat{\delta}) = \mathbf{Z}^T \hat{\mathbf{S}}^{-1} \mathbf{Z}$$

where $\mathbf{Z} = (Z_1, Z_2, \dots, Z_k)^T$, and

$$Z_j = \frac{1}{\sqrt{n}} (\mathcal{O}_{j,X} - e_{j,X}) |_{(j=1,2,\dots,k)},$$

and $\mathcal{O}_{j,X}$ reflect the total number of failures that have been noticed throughout these times. Due to Bagdonavičius et al. [1], Bagdonavičius and Nikulin [2], Bagdonavičius and Nikulin [3], Nikulin [4], Nikulin [5], Nikulin [6], and Rao and Robson [7], the test statistic is expressed as:

$$Y_{n,r-1,\varepsilon}^2(\hat{\delta}) = \sum_{j=1}^k \frac{1}{\mathcal{O}_{j,X}} (\mathcal{O}_{j,X} - e_{j,X})^2 + V_{W,G},$$

where $V_{W,G}$ and many other details are given in Bagdonavičius et al. [1], Bagdonavičius and Nikulin [2], Bagdonavičius and Nikulin [3], Nikulin [4], Nikulin [5], Nikulin [6], and Rao and Robson [7]. We compute each component of the $Y_{n,r-1,\varepsilon}^2(\hat{\delta})$ statistic for the OPPE model. The statistic $Y_{n,r-1,\varepsilon}^2(\hat{\delta})$ has a chi-square limit distribution, and its degree of freedom is $df = \text{rank}(\mathbf{S}) = \text{trace}(\mathbf{S}^{-1}\mathbf{S})$ where $df = k$. The estimated significance threshold is rejected if $Y_{n,r-1,\varepsilon}^2(\hat{\delta}) > \chi_{\varepsilon}^2(df)$ (where $\chi_{\varepsilon}^2(df)$ is the quantile of chi-square with $r = df = \text{rank}(\mathbf{S})$), then the approximate significance level ε is rejected Hypothesis, where $\hat{\mathbf{P}}_{I_j}$ is the Fisher information.

4.1. Uncensored Simulation Study under the NKRR Statistics $\mathbf{Y}^2(\varepsilon; \hat{\delta})$

We conducted an extensive study using numerical simulation to confirm the arguments of this work. Therefore, we generated the N statistics of 15,000 simulated samples with sizes of $n = 25$, $n = 50$, $n = 150$, $n = 350$, and $n = 600$, in order to test the null hypothesis H_0 that the sample belongs to the OPPE model. We determine the average of the non-rejection numbers

of the null hypothesis $Y^2(\varepsilon; \hat{\delta}) \leq \chi^2_\varepsilon(b-1)$ for various theoretical levels ($\varepsilon = 0.01, 0.02, 0.05, 0.1$). The corresponding empirical and theoretical levels are represented in Table 12. It can be seen that the calculated empirical level value is very close to its corresponding theoretical level value. Therefore, we conclude that the recommended test is very suitable for the OPPE distribution.

Table 12. Empirical levels and corresponding theoretical levels ($\varepsilon = 0.01, 0.02, 0.05, 0.1$) and $N = 15,000$.

$n \downarrow \& \varepsilon \rightarrow$	$\varepsilon_1 = 0.01$	$\varepsilon_2 = 0.02$	$\varepsilon_3 = 0.05$	$\varepsilon_4 = 0.1$
$n_1 = 25$	0.9941	0.9829	0.9522	0.9031
$n_2 = 50$	0.9936	0.9820	0.9513	0.9024
$n_3 = 150$	0.9922	0.9812	0.9510	0.9012
$n_5 = 300$	0.9909	0.9807	0.9507	0.9008
$n_6 = 600$	0.9905	0.9804	0.9503	0.9004

4.2. Uncensored Applications under the NKRR Statistics $Y^2(\varepsilon; \hat{\delta})$

4.2.1. Strengths of Glass Fibers

In this example, 100 carbon fiber fracture stresses (in Gba) are included in this data set given by Nicholase and Padgett [16]. Using the BB algorithm, we can get the MLE value of the parameter δ , assuming that our OPPE model can fit the strength data of 1.5 cm glass fiber: $\hat{\delta} = 2.14068$. We can compute and provide the Fisher Inf-Mx as follows using the value:

$$I(\hat{\delta}) = 1.648771.$$

The critical values for the NKRR statistical test were: $Y^2(\varepsilon; \hat{\delta}) = 11.855642$ and $\chi^2_{0.05}(6) = 12.59159$, so the OPPE distribution can effectively simulate and model the 1.5 cm glass fiber data.

4.2.2. Heat Exchanger Tube Crack

The crack data, which includes examinations carried out at eight chosen intervals until fractures emerged in 167 similar turbine parts, was acquired from the book by Meeker and Escobar [17].

Time of inspection	186	606	902	1077	1209	1377	1592	1932
Number of fans found to have cracks	5	16	12	18	18	2	6	17

We test the null hypothesis that these data are modified by our OPPE distribution using previously acquired NKRR statistics. We calculate the MLE $\hat{\delta} = 2.475019$ using R programming and the BB method (see Ravi (2009)). The estimated Fisher Inf-Mx at that time is:

$$I(\hat{\delta}) = 1.592466$$

So, we arrive with $Y^2(\varepsilon; \hat{\delta}) = 19.84927$ as the answer. The crucial value for significance level $\varepsilon = 0.05$ and $\chi^2_{0.01}(12) = 21.02607$. This model's NKRR statistic ($Y^2(\varepsilon; \hat{\delta})$) is less than the essential value, allowing us to conclude that the data correctly fit the OPPE model.

4.3. Censored Simulation Study under the NKRR Statistics $Y^2_{n,r-1,\varepsilon}(\hat{\delta})$

In the censored simulation study under the NKRR statistics Y^2 , it is expected that the produced sample ($N = 14000$) is censored at 25% with $df = 5$. We determine the average value of the non-rejection numbers of the null hypothesis for various theoretical levels ($\varepsilon = 0.01, 0.02, 0.05, 0.1$), where $Y^2_{n,r-1,\varepsilon}(\hat{\delta}) \leq \chi^2_\varepsilon(r-1)$. The matching theoretical and empirical levels are presented in Table 13, demonstrating how closely the determined empirical level value resembles the associated theoretical level value. Therefore, we draw the conclusion that the custom test is ideally suited to the OPPE model. These findings lead us to the conclusion that the theoretical level of the chi-square distribution on degrees

of freedom corresponds to the empirical significance level of the $Y_{n,r-1,\varepsilon}^2(\hat{\delta})$ statistics at which it is statistically significant. It may be inferred from this that the suggested test can accurately fit the censored data from the OPPE distribution. Returning to the defects of the NKRR basic test, the statistical literature contains many statistical hypothesis tests that fit the complete data, and for this reason, the original test has many competing tests in this field, and the fact is that the NKRR basic test actually has many alternatives. As for the modified NKRR test, it is considered an individual statistical test of its kind in the field of statistical tests due to its importance and the nature of the data that it can deal with, which are the data subject to censorship.

Table 13. Empirical levels and corresponding theoretical levels ($\varepsilon = 0.01, 0.02, 0.05, 0.1$) and $N = 14,000$.

$n \downarrow \&\varepsilon \rightarrow$	$\varepsilon_1 = 0.01$	$\varepsilon_2 = 0.02$	$\varepsilon_3 = 0.05$	$\varepsilon_4 = 0.1$
$n_1 = 25$	0.9930	0.9829	0.9533	0.9025
$n_2 = 50$	0.9925	0.9818	0.9524	0.9014
$n_3 = 150$	0.9912	0.9812	0.9516	0.9009
$n_5 = 300$	0.9906	0.9807	0.9508	0.9004
$n_6 = 600$	0.9904	0.9803	0.9502	0.9001

4.4. Censored Applications under the NKRR Statistics $Y_{n,x-1,\varepsilon}^2(\hat{\delta})$

4.4.1. Censored Lung Cancer Data Set

The lung cancer data provided by Loprinzi et al. [18] from the North Central Cancer Treatment Group investigated the survival in patients with advanced lung cancer ($n = 228$ and censored items = 63), and their performance scores rate how well the patient can do typical daily activities. If we assume that the data are distributed according to the OPPE distribution, we can estimate the vector parameter $\hat{\delta}$ by applying the maximum likelihood estimation approach as: $\hat{\delta} = 3.004756$. As a number of classes, we employ $df = 8$. The following is how the test statistic $Y_{n,r-1,\varepsilon}^2(\hat{\delta})$ items are presented:

$a_{j,b}(\hat{X})$	92.138	171.694	216.037	283.012	355.086	456.277	685.261	1022.3174
$\mathcal{O}_{j,X}$	29	30	35	31	32	25	28	18
$e_{j,X}$	6.55438	6.55438	6.55438	6.55438	6.55438	6.55438	6.55438	6.55438

The estimated matrix $\hat{P}_{lj}(X)$ and the estimated information (E-Inf) $\mathbf{I}(\hat{\delta})$ via Fisher are as follows:

$\hat{P}_{lj}(X)$	−0.6827	−0.4368	0.7015	0.7684	0.5094	0.3007	0.8769	−0.2884
-------------------	---------	---------	--------	--------	--------	--------	--------	---------

and

$$\mathbf{I}(\hat{\delta}) = 3.0134022.$$

The chi-squared test has a critical value of $\chi_{0.05}^2(df = 8) = 15.50731$. We find that the estimated statistic for the suggested test is $Y_{n,r-1,\varepsilon}^2(\hat{\delta}) = 14.616535$ using the earlier findings. We may state that our hypothesis H_0 is accepted because the tabulated value of the $Y_{n,r-1,\varepsilon}^2(\hat{\delta})$ statistic is higher than the computed value. This leads us to the conclusion that there is a 5% chance that the lung cancer data will deviate from the OPPE distribution.

4.4.2. Censored Capacitor Data Reliability Data Set

A set of data for basic reliability assessments data on glass capacitor longevity as a function of voltage and operating temperature from a factorial experiment (see Meeker and Escobar [17]). Each combination of temperature and voltage has eight capacitors. Testing was stopped after the fourth failure at each combination, $n = 64$ and censored items = 32. The parameter vector's maximum likelihood estimator $\hat{\delta}$, assuming that the

data are distributed according to the OPPE distribution, is: $\hat{\delta} = 1.680225$. We pick a few classes with $df = 8$. The components of the statistical test $Y_{n,r-1,\epsilon}^2(\hat{\delta})$ are as follows:

$a_{j,b}(\hat{X})$	346.1573	469.502	587.113	679.017	1078.834	1089.109	1102.167	1106.444
$\mathcal{O}_{j,X}$	11	15	6	10	6	5	6	5
$e_{j,X}$	6.91042	6.91042	6.91042	6.91042	6.91042	6.91042	6.91042	6.91042

The estimated matrix $\hat{P}_{lj}(X)$ and Fisher's E-Inf $I(\hat{\delta})$ are:

$\hat{P}_{lj}(X)$	0.48965	−0.73465	−0.40012	0.29784	−0.29467	−0.95113	0.60378	0.28564
-------------------	---------	----------	----------	---------	----------	----------	---------	---------

and

$$I(\hat{\delta}) = 4.1675894$$

The value of the statistical test $Y_{n,r-1,\epsilon}^2(\hat{\delta}) = 13.84577$ is then assessed. $\chi_{0.05}^2(8) = 15.50731 > Y_{n,r-1,\epsilon}^2(\hat{\delta})$ is the crucial value. We may conclude that the OPPE model is used to update the life statistics for glass capacitors.

5. Conclusions

The one parameter Poisson-exponential (OPPE) model, a new adaptable variation of the exponentiated exponential model, is introduced and studied in this article. Six well-known estimation techniques are investigated, discussed, and used: Cramer–von Mises, ordinary least square, L-moments, maximum likelihood, Kolmogorov, and weighted-least square. The Bayesian estimation under the squared error loss function was also devised and investigated. The use of two real data and two simulated data allows for a complete evaluation of all available techniques. The utility and adaptability of the OPPE distribution are illustrated using two real data sets. The new model outperforms many other competing models in modeling relief times and survival times, according to the Akaike Information Criterion, Consistent Akaike Information Criterion, Hannan–Quinn Information Criterion, Bayesian Information Criterion, Cramér–von Mises, and Anderson–Darling statistics. However, there are more results the reader can read by examining the paper and the applications, the following results can be specifically highlighted:

- The maximum likelihood method is still the most effective and trustworthy of the remaining classic approaches. Both the Bayesian technique and the Maximum Likelihood method are recommended for statistical modeling and applications.
- Under different loss functions, the Bayesian estimation is provided. Three loss functions, the generalized quadratic, the Linex, and the entropy, are used to produce the Bayesian estimators, and many useful details are provided.
- All of the provided estimation methods have been evaluated through simulation tests with particular parameter and controls (these simulation studies are all stated in the paper at the appropriate places).
- The BB algorithm for process estimation under censored samples is used to compare the Bayesian technique and the censored maximum likelihood method.
- It is shown in detail how the NKRR statistic is created for the OPPE model in the unfiltered instance. The results of a simulation research show that the OPPE model and the NKRR test are a good fit.
- A simulation study for evaluating statistics is described, along with the construction of the Bagdonavičius and Nikulin statistic for the novel model under the censored situation. Two real data applications are examined in a censored scenario; the first data are reliability information on capacitors, and the second data are information about lung cancer (medical data). We deduce from these applications that the suggested technique can successfully fit censored data from the OPPE distribution.

- For the uncensored strengths of glass fibers data, the critical values for the NKRR statistical test were: $Y^2(\varepsilon; \hat{\delta}) = 11.855642$ and $\chi_{0.05}^2(6) = 12.59159$, so the OPPE distribution can effectively simulate and model the uncensored 1.5 cm glass fiber data.
- For the uncensored heat exchanger tube crack data, the critical values for the NKRR statistical test were: $Y^2(\varepsilon; \hat{\delta}) = 19.84927$ and $\chi_{0.05}^2(6) = 12.59159$, so the OPPE distribution can effectively simulate and model the uncensored heat exchanger tube crack data.
- For the censored lung cancer data set: the value of the statistical test $Y_{n,r-1,\varepsilon}^2(\hat{\delta}) = 13.84577$, where $\chi_{0.05}^2(8) = 15.50731 > Y_{n,r-1,\varepsilon}^2(\hat{\delta}) = 14.616535$. Then, we conclude that the OPPE model can be used for modeling the censored lung cancer data set.
- For censored capacitor data reliability data set: the value of the statistical test $Y_{n,r-1,\varepsilon}^2(\hat{\delta}) = 13.84577$, where $\chi_{0.05}^2(8) = 15.50731 > Y_{n,r-1,\varepsilon}^2(\hat{\delta}) = 13.84577$. Then, we conclude that the OPPE model can be used for modeling the censored capacitor data reliability data set.
- It is worth noting that NKRR basic test, which is supported by complete data, is the most popular test in the last ten years. This is because it fits the complete truth data, and this is the case in most practical and applied cases in various fields. What distinguishes this statistical test also is the availability of ready-made statistical packages on the R program, whether for simulation or applications on actual data.
- However, the practical and experimental reality in many fields (such as the medical, chemical, engineering, etc.) necessitates that researchers deal with practical experiments that produce controlled censored data. This type of data, of course, needs certain tests dedicated to statistical dealing with it in the problem of distributional validation. The NKRR basic test is not the optimal choice in these cases. Hence, and based on this dilemma, was the primary and most important motive that prompted many researchers to think about introducing a new statistical test that fits the censored data. This new NKRR test, of course, is a modified test from the NKRR original test.
- According to the nature of the procedures of the two tests, both tests are not suitable for working with a type of data. For example, one of the drawbacks of the original test is that it is not suitable for dealing with censored data, it is only for complete data and cannot be applied to censored data. Therefore, if we are dealing with complete data, then this test may be a strong candidate for performing statistical hypothesis tests.
- Moreover, a disadvantage of the modified test is that it is only suitable for dealing with censored data, it is not intended for complete data and can only be applied to censored data. Therefore, if we are dealing with censored data, this test will undoubtedly be a strong candidate for statistical hypothesis tests, because it is intended for this type of data.
- Returning to the defects of the NKRR basic test, the statistical literature contains many statistical hypothesis tests that fit the complete data, and for this reason, the original test has many competing tests in this field, and the fact is that the NKRR basic test actually has many alternatives.
- As for the modified NKRR test, it is considered an individual statistical test of its kind in the field of statistical tests due to its importance and the nature of the data that it can deal with, which are the data subject to censorship.

Author Contributions: W.E.: validation, writing the original draft preparation, conceptualization, data curation, formal analysis, software. Y.T.: funding acquisition methodology, formal analysis, conceptualization, software. H.G.: validation, formal analysis, data duration, conceptualization, software. T.H.: validation, formal analysis, data duration, conceptualization, formal analysis, software. A.H.: validation, data duration, conceptualization, software. M.M.A.: review and editing, conceptualization, formal analysis, supervision. H.M.Y.: review and editing, software, validation, writing the original draft preparation, conceptualization, supervision. M.I.: review and editing, formal analysis, software, validation, writing the original draft preparation, conceptualization. All authors have read and agreed to the published version of the manuscript.

Funding: The study was funded by Researchers Supporting Project number (RSP2023R488), King Saud University, Riyadh, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset can be provided upon requested.

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

Abbreviations

CDF	Cumulative distribution function
OPPE	One parameter Poisson-exponential
SrF	Survival function
ADT	Anderson-Darling test
CVME	Cramér–von Mises estimation
MLEs	Maximum likelihood estimations
HRF	Hazard rate function
Inf-Mx	Information matrix
E- Inf	Estimated information
OLSQE	Ordinary least squares estimation
KE	Kolmogorov estimation
WLSQE	Weighted least squares estimation
MSE	Mean square error
$\ell(\delta)$	Log-likelihood
CVMC	Cramér–von Mises criterion
df	Degrees of freedom
P–P	Probability-probability
TTT	Total time in test
Q–Q	Quantile-quantile
NKRR	Nikulin–Rao–Robson

References

1. Bagdonavičius, V.B.; Levulienė, R.J.; Nikulin, M.S. Chi-Squared Goodness-of-Fit Tests for Parametric Accelerated Failure Time Models. *Commun. Stat.-Theory Methods* **2013**, *42*, 2768–2785.
2. Bagdonavičius, V.; Nikulin, M. Chi-squared goodness-of-fit test for right censored data. *Int. J. Appl. Math. Stat.* **2011**, *24*, 30–50.
3. Bagdonavičius, V.; Nikulin, M. Chi-squared tests for general composite hypotheses from censored samples. *Comptes Rendus Math.* **2011**, *349*, 219–223.
4. Nikulin, M.S. Chi-squared test for normality. In *Proceedings of the International Vilnius Conference on Probability Theory and Mathematical Statistics*; Vilnius University: Vilnius, Lithuania, 1973; Volume 2, pp. 119–122.
5. Nikulin, M.S. Chi-squared test for continuous distributions with shift and scale parameter. *Theory Probab. Its Appl.* **1973**, *18*, 559–568. [[CrossRef](#)]
6. Nikulin, M.S. On a Chi-squared test for continuous distributions. *Theory Probab. Its Appl.* **1973**, *19*, 638–639.
7. Rao, K.C.; Robson, D.S. A Chi-Square Statistic for Goodness-of-Fit Tests Within the Exponential Family. *Commun. Stat.-Simul. Comput.* **1974**, *3*, 1139–1153.
8. Ibrahim, M.; Hamedani, G.G.; Butt, N.S.; Yousof, H.M. Expanding the Nadarajah Haghighi Model: Copula, Censored and Uncensored Validation, Characterizations and Applications. *Pak. J. Stat. Oper. Res.* **2022**, *18*, 537–553. [[CrossRef](#)]

9. Yadav, A.S.; Shukla, S.; Goual, H.; Saha, M.; Yousof, H.M. Validation of xgamma exponential model via Nikulin-Rao-Robson goodness-of-fit test under complete and censored sample with different methods of estimation. *Stat. Optim. Inf. Comput.* **2022**, *10*, 457–483.
10. Gross, J.; Clark, V.A. *Survival Distributions: Reliability Applications in the Biometrical Sciences*; John Wiley: New York, NY, USA, 1975.
11. Bjerkedal, T. Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Am. J. Epidemiol.* **1960**, *72*, 130–148.
12. Muñoz-Gil, G.; Volpe, G.; Garcia-March, M.A.; Aghion, E.; Argun, A.; Hong, C.B.; Manzo, C. Objective comparison of methods to decode anomalous diffusion. *Nat. Commun.* **2021**, *12*, 6253. [\[CrossRef\]](#)
13. Pitman, E. The closest estimates of statistical parameter. In *Mathematical Proceedings of the Cambridge Philosophical Society*; Cambridge University Press: Cambridge, UK, 1937; Volume 33.
14. Fuller, W.A. Closeness of estimators. In *Encyclopedia of Statistical Sciences*; Wiley: New York, NY, USA, 1982; Volume 2.
15. Jozani, M.J.; Davies, K.F.; Balakrishnan, N. Pitman closeness results concerning ranked set sampling. *Stat. Probab. Lett.* **2012**, *82*, 2260–2269. [\[CrossRef\]](#)
16. Nicholase, M.D.; Padgett, W.J. A bootstrap control chart for Weibull percentiles. *Qual. Reliab. Eng. Int.* **2006**, *22*, 141–151. [\[CrossRef\]](#)
17. Meeker, W.Q.; Escobar, L.A.; Lu, C.J. Accelerated degradation tests: Modeling and analysis. *Technometrics* **1998**, *40*, 89–99. [\[CrossRef\]](#)
18. Loprinzi, C.L.; ALaurie, J.; Wieand, H.S.; EKrook, J.; Novotny, P.J.; Kugler, J.W.; Bartel, J.; Law, M.; Bateman, M.; EKlatt, N. Prospective evaluation of prognostic variables from patient-completed questionnaires. North Central Cancer Treatment Group. *J. Clin. Oncol.* **1994**, *12*, 601–607. [\[CrossRef\]](#)

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.