mathematics

MDPI

Article

Discriminating among Generalized Exponential, Weighted Exponential and Weibull Distributions

Ruizheng Niu [1], Weizhong Tian [2,*] and Yunchu Zhang [3]

[1] School of Science, Xi'an University of Technology, Xi'an 710048, China; 18329932502@163.com
[2] College of Big Data and Internet, Shenzhen Technology University, Shenzhen 518118, China
[3] College of New Material and New Energy, Shenzhen Technology University, Shenzhen 518118, China; 15965704668@163.com
* Correspondence: tianweizhong@sztu.edu.cn

Abstract: In this paper, we consider the problem of discriminating among three different positively skewed lifetime distributions, namely the generalized exponential distribution, the weighted exponential distribution, and the Weibull distribution. All of these distributions have been used quite effectively to analyze positively skewed lifetime data. We use the methods of the ratio of maximized likelihood, the minimum Kolmogorov distance, and the sequential probability ratio test to discriminate among these three distributions. The probability of correct selection is considered for each hypothesis based on several scenarios with Monte Carlo simulation. Real data applications are studied to illustrate the effectiveness of these proposed methods.

Keywords: generalized exponential distribution; weighted exponential distribution; Weibull distribution; model selection; probability of correct selection; sequential probability ratio test

MSC: 62E99; 62L10; 62F03; 62F07

Citation: Niu, R.; Tian, W.; Zhang, Y. Discriminating among Generalized Exponential, Weighted Exponential and Weibull Distributions. *Mathematics* **2023**, *11*, 3847. https://doi.org/10.3390/math11183847

Academic Editor: Christophe Chesneau

Received: 13 August 2023
Revised: 3 September 2023
Accepted: 5 September 2023
Published: 8 September 2023

1. Introduction

The problem of testing whether some given data come from one of two or more distributions is quite common in the statistical literature, see Atkinson [1,2], Chen [3], Chambers and Cox [4], Cox [5,6], and Dyer [7]. Several methods have been used on the problem of discriminating among different distributions, such as the ratio of maximized likelihood (RML), the minimum Kolmogorov distance (MKD), and the sequential probability ratio test (SPRT). In this paper, we discuss the methods of discriminating among the Weibull (WB) distribution, the generalized exponential (GE) distribution, and the weighted exponential (WE) distribution.

The WB distribution was introduced by Weibull [8] to analyze the skewed lifetime data, and has been widely applied in reliability engineering, failure analysis, and radar systems to model the dispersion of the received signal's level produced by some types of clutter. If x follows the WB distribution with the shape parameter β and scale parameter ξ, respectively, then it has the cumulative distribution function (cdf) and the probability density function (pdf) as follows

$$F_{WB}(x;\beta,\xi) = 1 - e^{-(\xi x)^{\beta}}, x > 0, \quad \text{and} \quad \beta,\xi > 0,$$

and

$$f_{WB}(x;\beta,\xi) = \beta\xi^{\beta}x^{\beta-1}e^{-(\xi x)^{\beta}}, x > 0, \quad \text{and} \quad \beta,\xi > 0,$$

respectively, and is denoted by $x \sim WB(\beta,\xi)$.

The GE distribution was proposed and studied quite extensively by Gupta and Kundu [9], and has the cdf

$$F_{GE}(x;\gamma,\theta) = (1 - e^{-\theta x})^{\gamma}, x > 0,$$

and the pdf

$$f_{GE}(x;\gamma,\theta) = \gamma\theta(1 - e^{-\theta x})^{\gamma-1}e^{-\theta x}, x > 0,$$

where $\gamma > 0, \theta > 0$ are shape and scale parameters, respectively, and it is denoted by $x \sim GE(\gamma,\theta)$. Moreover, the pdf of the distribution $GE(\gamma,\theta)$ strictly decreases when $\gamma \leq 1$, and has a unimodal shape for $\gamma > 1$. It is clear that the pdf of the $GE(\gamma,\theta)$ distribution is always right-skewed and can be used quite effectively to analyze skewed data sets.

Recently, Gupta and Kundu [10] introduced a new class of the WE distribution as a generalization of the standard exponential distribution, which has the pdf

$$f_{WE}(x;\eta,\lambda) = \frac{\eta+1}{\eta}\lambda e^{-\lambda x}(1 - e^{-\eta\lambda x}), x > 0,$$

and the cdf

$$F_{WE}(x;\eta,\lambda) = 1 + \frac{1}{\eta}e^{-\lambda(\eta+1)x} - \frac{\eta+1}{\eta}e^{-\lambda x}, x > 0,$$

where $\eta > 0$, $\lambda > 0$ are the shape and scale parameters, respectively, and it is denoted as $x \sim WE(\eta,\lambda)$. Furthermore, they showed that the WE distribution possesses some good properties and can be used as a good fit for survival time data compared to other popular distributions such as the gamma, the WB, or the GE distribution. For example, Makhdoom [11] investigated the statistical inference for reliability and stress strength for the WE distribution. Dey et al. [12] investigated various properties and methods of the estimation of the WE distribution. Tian and Yang [13] studied a change-point problem of the WE distribution based on the likelihood ratio test, modified information criterion, and Schwarz information criterion.

As we know, the WE distribution can be used as an alternative to the WB or GE distributions, and these distributions have several interesting properties. For certain values of shape and scale parameters, it is interesting to note that the shapes of these three distributions of cdfs and pdfs are very similar, but the hazard functions are completely different, which can be seen in Figures 1 and 2. These suggest that even if the distributions are very close in a sense of a certain distributional characteristic, they may be quite different concerning other characteristics.

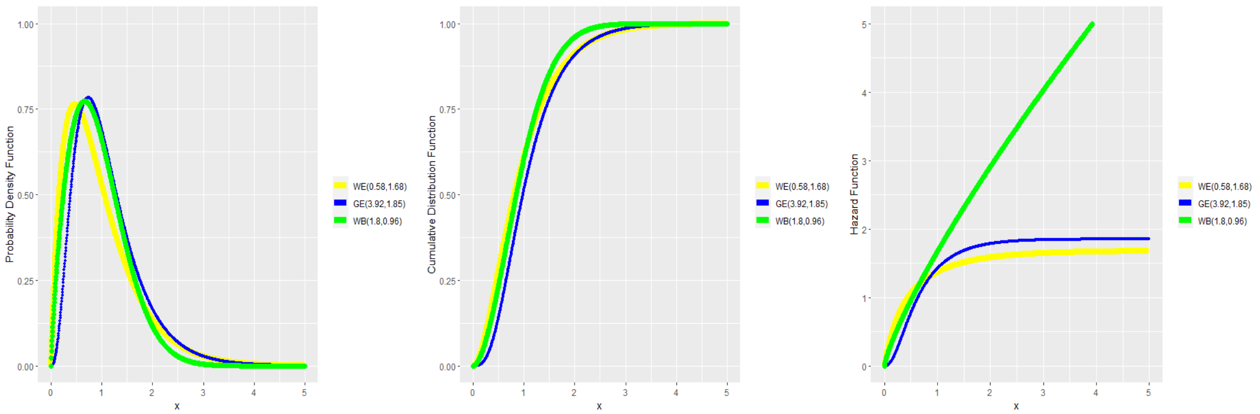

Figure 1. The pdfs, cdfs, and hazard functions of GE(3.92, 1.85), WE(0.58, 1.68), and WB(1.80, 0.96).

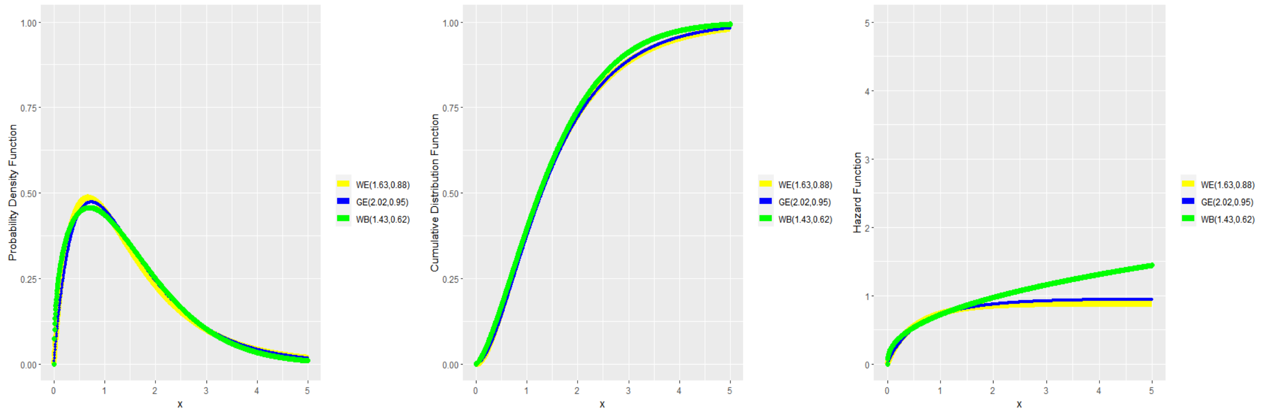

Figure 2. The pdfs, cdfs, and hazard functions of GE(2.02, 0.95), WE(1.63, 0.88), and WB(1.43, 0.62).

In fact, many scholars have studied the discrimination related to the WB distribution or the GE distribution. Gupta and Kundu [14] discussed the closeness of the gamma and the GE distributions using the RML and the MKD methods. Gupta and Kundu [15] considered the RML and the MKD methods to study the discriminating problem of the WB and the GE distributions. Pakyari [16] presented diagnostic tools based on the RML and the MKD methods to discriminate between the GE, the geometric extreme exponential, and the WB models. Raqab [17] considered the RML procedure to discriminate between the generalized Rayleigh and the WB distributions. Elsherpieny et al. [18] studied the RML procedure to discriminate between the gamma and the log-logistic distributions in the case of progressive type-II censoring. Ahmad et al. [19] used the MKD method to discriminate between the generalized Rayleigh and the WB distributions. Raqab et al. [20] studied the discriminating problem of the WB, the log-normal, and the log-logistic distributions based on the RML and the MKD methods. Vaidyanathan and Varghese [21] applied the RML and the MKD methods to discriminate between the exponential and the Lindley distributions based on a given random sample of observations, and this has been widely used in the fields of biology and engineering. Recently, Paul et al. [22] developed the SPRT methodology to discriminate between any two of the log-normal, the WB, and the log-logistic distributions as well as to discriminate among these distributions, which allowed the use of a less than average sample size without sacrificing the probability of correct selection (PCS).

As far as we know, statisticians have not yet studied the problem of discriminating among these distributions. Thus, it is significant to study this. The paper is organized as follows. Three different methods to discriminate among the GE, WE, and WB distributions are investigated in Section 2. Simulation studies are considered in Section 3. The real data set is analyzed in Section 4. Some concluding remarks are presented in Section 5.

2. Methodology

In this section, we study three different methods, namely the RML, the MKD, and the SPRT, to discriminate among the GE, the WE, and the WB distributions.

As a random sample, $x_1, x_2, \cdots, x_n$ is supposed to belong to one of the WE, GE, or WB distributions. Our interest lies in determining the distribution from which the random sample is observed. This problem can be formulated as the hypothesis-testing problem in the following

$$H_1 : x_1, x_2, \cdots \sim WE(\eta, \lambda) \text{ vs. } H_2 : x_1, x_2, \cdots \sim GE(\gamma, \theta) \text{ vs. } H_3 : x_1, x_2, \cdots \sim WB(\beta, \xi).$$

Under H_1, the log-likelihood function is given by

$$\log L_{WE} = n\left[\log\left(\frac{(\eta+1)\lambda}{\eta}\right) - \lambda\overline{X}\right] + \sum_{i=1}^{n} \log(1 - e^{-\eta\lambda x_i}), \tag{1}$$

where $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

The MLEs of $\hat{\eta}$ and $\hat{\lambda}$ are the solutions of the following equations after taking the first derivatives of (1) for η and λ,

$$\begin{aligned} \frac{\partial \log L_{WE}}{\partial \eta} &= \frac{n}{\eta(\eta+1)} - \sum_{i=1}^{n} \frac{\lambda x_i e^{-\eta\lambda x_i}}{1-e^{-\eta\lambda x_i}} = 0, \\ \frac{\partial \log L_{WE}}{\partial \lambda} &= \frac{n}{\lambda} - n\overline{X} + \sum_{i=1}^{n} \frac{\eta x_i e^{-\eta\lambda x_i}}{1-e^{-\eta\lambda x_i}} = 0. \end{aligned} \tag{2}$$

Under H_2, the log-likelihood function is given by

$$\log L_{GE} = n[\log(\gamma\theta) - \theta\overline{X}] + (\gamma-1)\sum_{i=1}^{n}\log(1-e^{-\theta x_i}). \tag{3}$$

Similarly, the MLEs of $\hat{\gamma}$ and $\hat{\theta}$ can be obtained by solving the following equations,

$$\begin{aligned} \frac{\partial \log L_{GE}}{\partial \gamma} &= \frac{n}{\gamma} - \sum_{i=1}^{n}\log(1-e^{-\theta x_i}) = 0, \\ \frac{\partial \log L_{GE}}{\partial \theta} &= \frac{n}{\theta} - n\overline{X} + (\gamma-1)\sum_{i=1}^{n}\frac{x_i e^{-\theta x_i}}{1-e^{-\theta x_i}} = 0. \end{aligned} \tag{4}$$

Under H_3, the log-likelihood function is given by

$$\log L_{WB} = n\log\beta + n\beta\log\xi + (\beta-1)\sum_{i=1}^{n}\log x_i - \sum_{i=1}^{n}(\xi x_i)^{\beta}. \tag{5}$$

Then, the MLEs of $\hat{\xi}$ and $\hat{\beta}$ can be obtained by solving the following equations,

$$\begin{aligned} \frac{\partial \log L_{WB}}{\partial \xi} &= \frac{n\beta}{\xi} - \beta\xi^{\beta-1}\sum_{i=1}^{n} {x_i}^{\beta} = 0, \\ \frac{\partial \log L_{WB}}{\partial \beta} &= \frac{n}{\beta} + n\log\xi + \sum_{i=1}^{n}\log x_i - \sum_{i=1}^{n}(\xi x_i)^{\beta}\log(\xi x_i) = 0. \end{aligned} \tag{6}$$

2.1. The Ratio of Maximized Likelihood Method

Cox [5,6] first proposed the RML test when discriminating between two separate models and derived the asymptotic distribution of the RML statistic. In the following, we introduce the procedure of the method of the RML.

The RMLs between two separate models can be defined as follows:

$$L_1 = \frac{L_{WE}(\hat{\eta},\hat{\lambda})}{L_{GE}(\hat{\gamma},\hat{\theta})}, \quad L_2 = \frac{L_{WB}(\hat{\beta},\hat{\xi})}{L_{WE}(\hat{\eta},\hat{\lambda})}, \quad L_3 = \frac{L_{WB}(\hat{\beta},\hat{\xi})}{L_{GE}(\hat{\gamma},\hat{\theta})}, \tag{7}$$

where $\hat{\eta}$, $\hat{\lambda}$, $\hat{\gamma}$, $\hat{\theta}$,$\hat{\beta}$, and $\hat{\xi}$ are calculated by (2), (4), and (6), respectively.

The logarithm of the above RMLs can be written as

$$
\begin{aligned}
T_n^1 &= \log L_1 \\
&= \log L_{WE}(\hat{\eta}, \hat{\lambda}) - \log L_{GE}(\hat{\gamma}, \hat{\theta}) \\
&= n\left[\log\left(\frac{(\hat{\eta}+1)\hat{\lambda}}{\hat{\eta}\hat{\gamma}\hat{\theta}}\right) - (\hat{\lambda}-\hat{\theta})\overline{X}\right] + \sum_{i=1}^{n}\log\left(\frac{1-e^{-\hat{\eta}\hat{\lambda}x_i}}{(1-e^{-\hat{\theta}x_i})^{\hat{\gamma}-1}}\right), \\
T_n^2 &= \log L_2 \\
&= \log L_{WB}(\hat{\beta}, \hat{\xi}) - \log L_{WE}(\hat{\eta}, \hat{\lambda}) \\
&= n\left[\log\left(\frac{\hat{\beta}\hat{\eta}\hat{\xi}^{\hat{\beta}}}{(\hat{\eta}+1)\hat{\lambda}}\right) + \hat{\lambda}\overline{X}\right] + \sum_{i=1}^{n}\log\left(\frac{x_i^{\hat{\beta}-1}}{1-e^{-\hat{\eta}\hat{\lambda}x_i}}\right) - \sum_{i=1}^{n}(\hat{\xi}x_i)^{\hat{\beta}}, \\
T_n^3 &= \log L_3 \\
&= \log L_{WB}(\hat{\beta}, \hat{\xi}) - \log L_{GE}(\hat{\gamma}, \hat{\theta}) \\
&= n\left[\log\left(\frac{\hat{\beta}\hat{\xi}^{\hat{\beta}}}{\hat{\gamma}\hat{\theta}}\right) + \hat{\theta}\overline{X}\right] + \sum_{i=1}^{n}\log\left(\frac{x_i^{\hat{\beta}-1}}{(1-e^{-\hat{\theta}x_i})^{\hat{\gamma}-1}}\right) - \sum_{i=1}^{n}(\hat{\xi}x_i)^{\hat{\beta}}.
\end{aligned}
\tag{8}
$$

Now we choose the WE distribution if $T_n^1 > 0$, $T_n^2 < 0$, the GE distribution if $T_n^1 < 0$, $T_n^3 < 0$, and the WB distribution if $T_n^2 > 0$, $T_n^3 > 0$.

2.2. The Minimum Kolmogorov Distance Method

The MKD test was originally employed to test the hypothesis that a completely random sample has come from a fully specified continuous distribution. The procedure of the MKD method is given as follows. Denote $\hat{F}_{WE}(\hat{\eta}, \hat{\lambda})$, $\hat{F}_{GE}(\hat{\gamma}, \hat{\theta})$, and $\hat{F}_{WB}(\hat{\beta}, \hat{\xi})$ as the cdfs calculated at the MLEs of the parameters of the WE, GE, and WB distributions, respectively, and let $\widetilde{F}(x)$ be the empirical distribution function calculated from the data. The KDs associated with the three models are given as, respectively,

$$
\begin{aligned}
KD_{WE} &= \sup_{-\infty<x<+\infty} |\hat{F}_{WE}(x, \hat{\eta}, \hat{\lambda}) - \widetilde{F}(x)|, \\
KD_{GE} &= \sup_{-\infty<x<+\infty} |\hat{F}_{GE}(x, \hat{\gamma}, \hat{\theta}) - \widetilde{F}(x)|, \\
KD_{WB} &= \sup_{-\infty<x<+\infty} |\hat{F}_{WB}(x, \hat{\beta}, \hat{\xi}) - \widetilde{F}(x)|.
\end{aligned}
\tag{9}
$$

The model with the minimum distance is then chosen as the winning model.

2.3. The SPRT Method

The SPRT analysis, first introduced by Wald [23], was originally developed as a more effective method of quality control during the Second World War. In this section, an SPRT procedure for discriminating the WE, the GE, and the WB is described.

As the data are observed sequentially, the goal is to identify the correct model by testing multiple hypotheses. The proposed discriminating method is based on the logarithm of the RMLs defined as:

$$
\begin{aligned}
\Lambda_n^1 &= \log\left(\frac{L_{GE}(x_i;\hat{\gamma}_n,\hat{\theta}_n)}{L_{WE}(x_i;\hat{\eta}_n,\hat{\lambda}_n)}\right) \\
&= \log L_{GE}(x_i;\hat{\gamma}_n,\hat{\theta}_n) - \log L_{WE}(x_i;\hat{\eta}_n,\hat{\lambda}_n) \\
&= n\left[\log\left(\frac{\hat{\eta}_n\hat{\gamma}_n\hat{\theta}_n}{(\hat{\eta}_n+1)\hat{\lambda}_n}\right) - (\hat{\theta}_n - \hat{\lambda}_n)\overline{X}\right] + \sum_{i=1}^{n}\log\left(\frac{(1-e^{-\hat{\theta}_n x_i})^{\hat{\gamma}_n-1}}{1-e^{-\hat{\eta}_n\hat{\lambda}_n x_i}}\right), \\
\Lambda_n^2 &= \log\left(\frac{L_{WE}(x_i;\hat{\eta}_n,\hat{\lambda}_n)}{L_{WB}(x_i;\hat{\beta}_n,\hat{\xi}_n)}\right) \\
&= \log L_{WE}(x_i;\hat{\eta}_n,\hat{\lambda}_n) - \log L_{WB}(x_i;\hat{\beta}_n,\hat{\xi}_n) \\
&= n\left[\log\left(\frac{(\hat{\eta}_n+1)\hat{\lambda}_n}{\hat{\beta}_n\hat{\eta}_n\hat{\xi}_n^{\hat{\beta}_n}}\right) - \hat{\lambda}_n\overline{X}\right] + \sum_{i=1}^{n}\log\left(\frac{1-e^{-\hat{\eta}_n\hat{\lambda}_n x_i}}{x_i^{\hat{\beta}_n-1}}\right) + \sum_{i=1}^{n}(\hat{\xi}_n x_i)^{\hat{\beta}_n}, \\
\Lambda_n^3 &= \log\left(\frac{L_{GE}(x_i;\hat{\gamma}_n,\hat{\theta}_n)}{L_{WB}(x_i;\hat{\beta}_n,\hat{\xi}_n)}\right) \\
&= \log L_{GE}(x_i;\hat{\gamma}_n,\hat{\theta}_n) - \log L_{WB}(x_i;\hat{\beta}_n,\hat{\xi}_n) \\
&= n\left[\log\left(\frac{\hat{\gamma}_n\hat{\theta}_n}{\hat{\beta}_n\hat{\xi}_n^{\hat{\beta}_n}}\right) - \hat{\theta}_n\overline{X}\right] + \sum_{i=1}^{n}\log\left(\frac{(1-e^{-\hat{\theta}_n x_i})^{\hat{\gamma}_n-1}}{x_i^{\hat{\beta}_n-1}}\right) + \sum_{i=1}^{n}(\hat{\xi}_n x_i)^{\hat{\beta}_n},
\end{aligned}
\tag{10}
$$

where $\hat{\eta}_n$, $\hat{\lambda}_n$, $\hat{\gamma}_n$, $\hat{\theta}_n$, $\hat{\beta}_n$, and $\hat{\xi}_n$ are MLEs of η, λ, γ, θ, β, and ξ, respectively.

A natural idea for multi-hypothesis testing is to select the hypothesis having the maximum likelihood. The SPRT procedure proposed here is based on Λ_n^1, Λ_n^2, and Λ_n^3 as defined in (10), respectively. The appropriate selection of stopping boundaries in (10) is crucial for ensuring a high PCS and savings in the sample size. Here, we choose $PCS \geq 1-\alpha$, where $\alpha \in (0,1)$ is the fixed level, and the SPRT stopping boundaries are $b_j = \ln(\frac{\alpha}{2}) = -a_j$, for $j = 1,2,3$. Let k $(1 < k \leq n)$ be a prefixed pilot sample size. The purely sequential procedure is discussed as follows:

Stage 1: Draw samples $x_1, x_2, \cdots, x_k$ and compute Λ_k^1, Λ_k^2, and Λ_k^3. Stop sampling if $\{\Lambda_k^1 \geq a_1$ and $\Lambda_k^3 \geq a_3\}$ or $\{\Lambda_k^1 \leq b_1$ and $\Lambda_k^2 \geq a_2\}$ or $\{\Lambda_k^2 \leq b_2$ and $\Lambda_k^3 \leq b_3\}$. Otherwise, proceed to the next stage.

Stage 2: Draw a new observation x_{k+1} independent of $x_1, x_2, \cdots, x_k$ and compute Λ_{k+1}^1, Λ_{k+2}^2, and Λ_{k+3}^3. Stop sampling if $\{\Lambda_{k+1}^1 \geq a_1$ and $\Lambda_{k+1}^3 \geq a_3\}$ or $\{\Lambda_{k+1}^1 \leq b_1$ and$\Lambda_{k+1}^2 \geq a_2\}$ or $\{\Lambda_{k+1}^2 \leq b_2$ and $\Lambda_{k+1}^3 \leq b_3\}$. Otherwise, proceed to draw one new observation. We continue sampling one observation at a time until

$$
t = \underset{n\geq k}{inf}\left\{\{\Lambda_n^1 \geq a_1 \cap \Lambda_n^3 \geq a_3\} \cup \{\Lambda_n^1 \leq b_1 \cap \Lambda_n^2 \geq a_2\} \cup \{\Lambda_n^2 \leq b_2 \cap \Lambda_n^3 \leq b_3\}\right\}. \tag{11}
$$

We select the GE distribution if $\{\Lambda_t^1 \geq a_1 \ \ and \ \ \Lambda_t^3 \geq a_3\}$, the WE distribution if $\{\Lambda_t^1 \leq b_1 \ \ and \ \ \Lambda_t^2 \geq a_2\}$, and the WB distribution if $\{\Lambda_t^2 \leq b_2 \ \ and \ \ \Lambda_t^3 \leq b_3\}$. In other words, the stopping time $T = min\{\tau_1, \tau_2, \tau_3\}$ where

$$
\tau_1 = \underset{n\geq k}{inf}\{\Lambda_n^1 \geq a_1, \Lambda_n^3 \geq a_3\}, \quad \tau_2 = \underset{n\geq k}{inf}\{\Lambda_n^1 \leq b_1, \Lambda_n^2 \geq a_2\}, \quad \tau_3 = \underset{n\geq k}{inf}\{\Lambda_n^2 \leq b_2, \Lambda_n^3 \leq b_3\}.
$$

2.4. Algorithm

According to the above procedure, the algorithm for each method can be summarized as follows:

(i) Generating a sample with sample size n follows the null hypothesis.
(ii) Using the sample to compute the corresponding MLE of the distribution follows H_1, H_2, H_3.
(iii) Compute the statistic of each method by using Equations (8), (9), and (10), respectively.
(iv) Select the correct distribution according to the criteria from each procedure.

3. Numerical Results

In this section, we performed some simulation experiments to compare different methods to discriminate among these three distributions. We used the RML and MKD methods first, then considered the SPRT procedure. All the computations were performed using R version 4.1.1 program [24]. To compare the performances of the different methods for different sample sizes and for different parameter values, we generated samples from the three different distributions and computed the PCS for each method based on 1000 replications.

We used the Monte Carlo simulation to generate data sets from each of the null distributions. In this case, we considered different sample sizes, namely $n = 20, 40, 60, 80$, 100, 200, and 500. For each of these combinations of the sample size, the shape parameter, and the scale parameter, we estimated the parameters of the three models by the MLEs procedure as described in (2), (4) and (6). To compare the performances of the different methods for different sample sizes and for different parameter values, we generated samples from the three different distributions and computed the PCS for each method based on 1000 replications. The details are explained below.

In the RML procedure, the RML statistic was calculated and the winning model was determined by the decision rule in Section 2.1. We repeated this procedure 1000 times and calculated the proportion of times the null distribution was chosen as the corresponding PCS. Moreover, we chose the values of the shape parameters that were greater than 1, and which had unimodal shapes of pdfs and had the significance of discrimination. For simplicity, we selected the same values of the scale parameters. Meanwhile, the corresponding PCSs for each hypothesis were calculated based on the procedure in Section 2.1.

I: The data come from the WE distribution. In this case, we set $\eta = 5.0, 5.2, 5.4, 5.8$ and $\lambda = 0.5, 1.0, 1.5$. We compute T_n^1 and T_n^2 as defined in (8) and the PCS_{WE}^{RML} can be written as follows,

$$PCS_{WE}^{RML} = P(T_n^1 > 0,\ T_n^2 < 0| \quad \text{data from the WE distribution}).$$

II: The data come from the GE distribution. We set $\gamma = 2.5, 2.7, 2.9, 3.3$ and $\theta = 0.5, 1, 1.5$, and the PCS_{GE}^{RML} can be obtained as follows,

$$PCS_{GE}^{RML} = P(T_n^1 < 0,\ T_n^3 < 0| \quad \text{data from the GE distribution}).$$

III: The data come from the WB distribution. We set the parameters $\beta = 2.5, 2.7, 2.9, 3.3$ and $\xi = 0.5, 1, 1.5$, and the PCS_{WB}^{RML} can be calculated as follows,

$$PCS_{WB}^{RML} = P(T_n^2 > 0,\ T_n^3 > 0| \quad \text{data from the WB distribution}).$$

Similarly, as in Pakyari [16] and Raqab et al. [20], we found the MLEs of parameters for the GE, WE, and WB distributions. The empirical distribution function was also calculated for each of the generated data sets. The model with the minimum distance was then chosen as the winning model. This procedure was repeated 1000 times and the corresponding PCS based on the MKD procedure, as described in Section 2.2, can be written as follows:

$$\begin{aligned}
PCS_{WE}^{MKD} &= P(KD_{WE} < KD_{GE},\ KD_{WB} > KD_{WE}| \quad \text{data from the WE distribution}),\\
PCS_{GE}^{MKD} &= P(KD_{WE} > KD_{GE},\ KD_{WB} > KD_{GE}| \quad \text{data from the GE distribution}),\\
PCS_{WB}^{MKD} &= P(KD_{WB} < KD_{WE},\ KD_{WB} < KD_{WE}| \quad \text{data from the WB distribution}).
\end{aligned}$$

Tables 1–3 give the PCS under the RML and the MKD procedures when the data come from different distributions. In Table 1, the PCS_{RML} is higher than PCS_{MKD}, but only slightly higher in some cases, which shows that the difference in the PCS choice between the two methods is not very large. At the same time, when the sample size is bigger than 100, the PCS can only reach 70%, which means it may take a certain sample size to distinguish among the three distributions when the data are generated from the WE distribution. In Table 2, the values of PCS for these two methods are improved as the shape parameter and scale parameter increase. However, the value of PCS_{RML} is significantly higher than that of the PCS_{MKD}. In Table 3, the value of PCS is close to 1 when the sample sizes exceed 200. In all three tables, we find that the PCS increases as the sample size and parameters improve, and the RML method is superior to the MKD procedure.

Table 1. The PCSs based on the Monte Carlo simulation when using data from the WE distribution.

(η, λ)	PCS	20	40	60	80	100	200	500
(5, 0.5)	RML	0.5072	0.5956	0.6446	0.6700	0.6976	0.7527	0.8570
	MKD	0.3992	0.4562	0.5012	0.5563	0.6112	0.7002	0.7882
(5.2, 0.5)	RML	0.5216	0.6102	0.6522	0.6862	0.6990	0.7642	0.8552
	MKD	0.4002	0.4577	0.5032	0.5575	0.6155	0.7156	0.7889
(5.4, 0.5)	RML	0.5288	0.6136	0.6544	0.6844	0.7114	0.7654	0.8628
	MKD	0.4132	0.4585	0.5045	0.5584	0.6225	0.7226	0.7995
(5.8, 0.5)	RML	0.5364	0.6264	0.6630	0.6950	0.7152	0.7770	0.8702
	MKD	0.4235	0.4598	0.5112	0.5598	0.6310	0.7250	0.8002
(5, 1)	RML	0.5152	0.6066	0.6486	0.6760	0.6966	0.7582	0.8588
	MKD	0.4005	0.4587	0.5022	0.5589	0.6136	0.7125	0.7902
(5.2, 1)	RML	0.5236	0.6192	0.6540	0.6872	0.6996	0.7666	0.8564
	MKD	0.4050	0.4592	0.5044	0.5591	0.6162	0.7163	0.7905
(5.4, 1)	RML	0.5340	0.6216	0.6594	0.6896	0.7132	0.7692	0.8666
	MKD	0.4138	0.4598	0.5199	0.5589	0.6255	0.7320	0.8002
(5.8, 1)	RML	0.5386	0.6348	0.6740	0.6972	0.7158	0.7802	0.8744
	MKD	0.4244	0.4608	0.5220	0.5623	0.6332	0.7365	0.8115
(5, 1.5)	RML	0.5176	0.6072	0.6584	0.6856	0.7030	0.7602	0.8596
	MKD	0.4010	0.4599	0.5056	0.5623	0.6188	0.7223	0.7956
(5.2, 1.5)	RML	0.5312	0.6224	0.6562	0.6918	0.7094	0.7690	0.8654
	MKD	0.4112	0.4602	0.5089	0.5633	0.6220	0.7228	0.7998
(5.4, 1.5)	RML	0.5422	0.6240	0.6622	0.6914	0.7164	0.7728	0.8710
	MKD	0.4189	0.4625	0.5232	0.5662	0.6305	0.7335	0.8156
(5.8, 1.5)	RML	0.5454	0.6358	0.6694	0.7006	0.7188	0.7830	0.8774
	MKD	0.4256	0.4668	0.5238	0.5671	0.6354	0.7399	0.8226

Table 2. The PCSs based on the Monte Carlo simulation when using data from the GE distribution.

(γ, θ)	PCS	20	40	60	80	100	200	500
(2.5, 0.5)	RML	0.4690	0.5460	0.6124	0.6612	0.7070	0.8168	0.9482
	MKD	0.3746	0.4156	0.4620	0.5042	0.5218	0.6410	0.8226
(2.7, 0.5)	RML	0.5056	0.6176	0.6698	0.7276	0.7706	0.8874	0.9808
	MKD	0.4098	0.4758	0.5040	0.5554	0.5934	0.7382	0.9170
(2.9, 0.5)	RML	0.5560	0.6578	0.7252	0.7774	0.8094	0.9200	0.9866
	MKD	0.4116	0.4889	0.5051	0.5561	0.5955	0.7392	0.9188
(3.3, 0.5)	RML	0.5930	0.7102	0.7728	0.8180	0.8512	0.9370	0.9932
	MKD	0.4225	0.4902	0.5066	0.5575	0.5965	0.7399	0.9192
(2.5, 1)	RML	0.4862	0.5596	0.6182	0.6648	0.7082	0.8192	0.9504
	MKD	0.3756	0.4174	0.4648	0.5052	0.5240	0.6468	0.8304
(2.7, 1)	RML	0.5138	0.6184	0.6706	0.7334	0.7788	0.8878	0.9812
	MKD	0.4105	0.4768	0.5063	0.5559	0.5996	0.7388	0.9220
(2.9, 1)	RML	0.5576	0.6620	0.7276	0.7812	0.8112	0.9212	0.9896
	MKD	0.4220	0.4992	0.5065	0.5602	0.5998	0.7401	0.9232
(3.3, 1)	RML	0.6006	0.7136	0.7766	0.8186	0.8532	0.9376	0.9936
	MKD	0.4236	0.4933	0.5074	0.5678	0.6002	0.7411	0.9235
(2.5, 1.5)	RML	0.4894	0.5604	0.6224	0.6696	0.7188	0.8254	0.9510
	MKD	0.3804	0.4238	0.4690	0.5070	0.5260	0.6514	0.8322
(2.7, 1.5)	RML	0.5154	0.6278	0.6790	0.7354	0.7830	0.8892	0.9842
	MKD	0.4188	0.4777	0.5077	0.5613	0.6005	0.7416	0.9235
(2.9, 1.5)	RML	0.5592	0.6716	0.7326	0.7848	0.8128	0.9238	0.9910
	MKD	0.4222	0.5009	0.5112	0.5622	0.6015	0.7554	0.9245
(3.3, 1.5)	RML	0.6098	0.7208	0.7820	0.8246	0.8536	0.9398	0.9954
	MKD	0.4288	0.4955	0.5142	0.5688	0.6116	0.7623	0.9255

Table 3. The PCSs based on the Monte Carlo simulation when using data from the WB distribution.

(γ, θ)	PCS	20	40	60	80	100	200	500
(2.5, 0.5)	RML	0.7356	0.8264	0.8894	0.9264	0.9458	0.9888	1
	MKD	0.4326	0.6158	0.7206	0.7882	0.8324	0.9418	0.9956
(2.7, 0.5)	RML	0.7660	0.8600	0.9058	0.9368	0.9532	0.9944	1
	MKD	0.4628	0.6406	0.7564	0.8106	0.8590	0.9576	0.9978
(2.9, 0.5)	RML	0.7698	0.8642	0.9214	0.9452	0.9682	0.9954	1
	MKD	0.4816	0.6566	0.7652	0.8250	0.8688	0.9630	0.9990
(3.3, 0.5)	RML	0.7946	0.8850	0.9348	0.9634	0.9782	0.9984	1
	MKD	0.5260	0.6880	0.8044	0.8604	0.9074	0.9782	0.9996
(2.5, 1)	RML	0.7406	0.8284	0.8968	0.9294	0.9480	0.9914	1
	MKD	0.4450	0.6214	0.7232	0.7912	0.8416	0.9442	0.9958
(2.7, 1)	RML	0.7628	0.8610	0.9084	0.9396	0.9588	0.9948	1
	MKD	0.4633	0.6423	0.7612	0.8226	0.8599	0.9662	0.9981
(2.9, 1)	RML	0.7746	0.8680	0.9228	0.9462	0.9696	0.9966	1
	MKD	0.4822	0.6589	0.7668	0.8288	0.8696	0.9676	0.9995
(3.3, 1)	RML	0.7996	0.8868	0.9350	0.9664	0.9784	0.9986	1
	MKD	0.5311	0.6885	0.8123	0.8698	0.9122	0.9881	0.9998
(2.5, 1.5)	RML	0.7468	0.8456	0.8990	0.9302	0.9522	0.9922	1
	MKD	0.4466	0.6234	0.7252	0.7930	0.8426	0.9464	0.9970
(2.7, 1.5)	RML	0.7672	0.8694	0.9088	0.9442	0.9604	0.9954	1
	MKD	0.4665	0.6466	0.7655	0.8239	0.8623	0.9702	0.9985
(2.9, 1.5)	RML	0.7778	0.8768	0.9266	0.9520	0.9714	0.9974	1
	MKD	0.4836	0.6612	0.7670	0.8295	0.8702	0.9702	0.9998
(3.3, 1.5)	RML	0.8002	0.8912	0.9384	0.9678	0.9804	0.9992	1
	MKD	0.5326	0.6892	0.8133	0.8706	0.9222	0.9905	1

Next, in the SPRT procedure, we computed Λ_t^1, Λ_t^2, and Λ_t^3 as defined in (10) and then found the best-fitting distribution based on the criterion given in Section 2.3. We considered the SPRT procedure for some prefixed values of $\alpha = 0.1, 0.2$. Meanwhile, we set $k = 10$ fixed sample sizes here and repeated the purely sequential procedure 1000 times. Using the same procedures as Paul et al. [22], the $PCSs$ are obtained in the following,

$$
\begin{aligned}
PCS_{WE}^{SPRT} &= P(\Lambda_t^1 \le b_1,\ \ \Lambda_t^2 \ge a_2 |\ \ \text{data from the WE distribution}),\\
PCS_{GE}^{SPRT} &= P(\Lambda_t^1 \ge a_1,\ \ \Lambda_t^3 \ge a_3 |\ \ \text{data from the GE distribution}),\\
PCS_{WB}^{SPRT} &= P(\Lambda_t^2 \le b_2,\ \ \Lambda_t^3 \le b_3 |\ \ \text{data from the WB distribution}).
\end{aligned}
$$

Tables 4–6 compare the performances of the fixed-sample procedure; here we pick the RML procedure to discriminate among these three distributions with the SPRT procedure. $E(t)$ and n_{RML} denote the estimated expected sample size for the SPRT procedure and the sample size of the RML procedure. n_{RML} is chosen such that PCS_{RML} is approximately equal to PCS_{SPRT} without sacrificing the PCS. The last columns of these tables report the savings in the average (expected) sample sizes as $\frac{n_{RML}-E(t)}{n_{RML}} \times 100\%$.

Table 4. Comparing the method of RML discrimination and the SPRT procedure when using data from the WE distribution.

True Distribution	Boundaries $\lvert b_j \rvert = a_j = \lvert \ln(\frac{\alpha}{2}) \rvert$	PCS_{SPRT}	PCS_{RML}	$E(t)$	n_{RML}	Savings %
WE (5, 0.5)	2.9957	0.9662	0.9605	521.33	800	34.83
	2.3026	0.9009	0.9025	483.66	700	30.91
WE (5.2, 0.5)	2.9957	0.9447	0.9436	506.66	750	32.45
	2.3026	0.9118	0.9156	452.22	660	34.63
WE (5.4, 0.5)	2.9957	0.9502	0.9500	485.99	690	29.57
	2.3026	0.9100	0.9164	421.20	610	30.95
WE (5.8, 0.5)	2.9957	0.9409	0.9438	458.66	640	28.33
	2.3026	0.9009	0.9066	396.88	560	29.13

Table 5. Comparing the method of RML discrimination and the SPRT procedure when using data from the GE distribution.

True Distribution	Boundaries $\lvert b_j \rvert = a_j = \lvert \ln(\frac{\alpha}{2}) \rvert$	PCS_{SPRT}	PCS_{RML}	$E(t)$	n_{RML}	Savings %
GE (2.5, 0.5)	2.9957 2.3026	0.9420 0.9002	0.9446 0.9056	452.22 411.32	780 660	42.02 37.68
GE (2.7, 0.5)	2.9957 2.3026	0.9500 0.9118	0.9588 0.9156	422.60 395.66	710 620	40.48 36.18
GE (2.9, 0.5)	2.9957 2.3026	0.9444 0.9103	0.9494 0.9111	400.22 378.60	650 580	38.43 34.72
GE (3.3, 0.5)	2.9957 2.3026	0.9599 0.9101	0.9556 0.9119	385.89 346.22	600 550	35.69 37.05

Table 6. Comparing the method of RML discrimination and the SPRT procedure when using data from the WB distribution.

True Distribution	Boundaries $\lvert b_j \rvert = a_j = \lvert \ln(\frac{\alpha}{2}) \rvert$	PCS_{SPRT}	PCS_{RML}	$E(t)$	n_{RML}	Savings %
WB (2.5, 0.5)	2.9957 2.3026	0.9433 0.9006	0.9450 0.9063	243.22 205.12	500 380	51.36 46.02
WB (2.7, 0.5)	2.9957 2.3026	0.9336 0.9112	0.9380 0.9156	228.62 186.23	450 340	49.20 45.23
WB (2.9, 0.5)	2.9957 2.3026	0.9503 0.9050	0.9551 0.9066	206.88 206.88	410 320	49.54 35.35
WB (3.3, 0.5)	2.9957 2.3026	0.9333 0.9034	0.9368 0.9067	206.88 126.33	380 290	45.56 56.44

From Tables 4–6, it is clear that as α increases, the value of n_{RML} and $E(t)$ decrease. With approximately equal PCSs, the SPRT methods require a much lower expected sample size than that of the corresponding RML procedure. Tables 4–6 illustrate that the SPRT procedure requires fewer observations to draw the same conclusion than the the fixed-sample procedure.

4. Data Analysis

In this section, we analyze a real data set and use the above methods to discriminate the WE, the GE and the WB distributions.

4.1. Malignant Melanoma Data

This data set is available in Andersen et al. [25] and R package "*timereg*" (see, Scheike and Zhang [26]), and contains the survival times (in years) of 205 patients after surgery for malignant melanoma (skin cancer) between the years 1962 and 1977 collected at the University Hospital of Odense, Denmark, by K.T. Drzewiecki. The data set can also be found in Appendix A.

We analyzed the data set based on the RML, the MKD, and the SPRT methods, and the results are summarized in Table 7. For the SPRT procedure, we set $a_j = 2.3026$, $b_j = -2.3026$, and j = 1, 2, 3. From the table, we can find that $T_n^2 = 23.4740 > 0$ and $T_n^3 = 16.8645 > 0$, which means the WB distribution is selected. Similarly, the MKDs for these three distributions are 0.1700, 0.1813, and 0.1127, respectively. Therefore, the MKD criterion also suggests choosing the WB model for this data set. Next, $\Lambda_t^2 = -2.3319 < -2.3026$ and $\Lambda_t^3 = -2.3102 < -2.3026$, which illustrates that the SPRT procedure also supports this fact in this case. However, it is remarkable that the average sample size $E(t)$ is only 106, which is much less than $n_F = 205$, giving 48.29% savings in the average sample size.

Table 7. Summary of the RML, the MKD, and the SPRT procedures and the goodness of fit test of the malignant melanoma data.

Model	GE	WE	WB
shape parameters' estimation scale parameters' estimation	$\hat{\gamma}$ = 2.3371 $\hat{\theta}$ = 0.2712	$\hat{\eta}$ = 2.1321 $\hat{\lambda}$ = 0.2152	$\hat{\beta}$ = 1.8895 $\hat{\xi}$ = 0.1521
T_n^1 T_n^2 T_n^3		−6.6095 23.4740 16.8645	
Λ_t^1 Λ_t^2 Λ_t^3		2.1162 −2.3319 −2.3102	
AIC BIC	1081.2700 1087.9160	1094.4870 1101.1330	1047.5400 1054.1860
MKD	0.1700	0.1813	0.1127

Figure 3 depicts the histogram of the malignant melanoma data, and pdfs of the WE, the GE, and the WB fitted models. From the fitted density functions, it appears that the WB distribution provides a better fit than the WE and the GE distributions in this case. Andersen et al. [25] suggested using a WB model for this data set and Paul et al. [22] used SPRT criteria to choose the WB distribution, which all confirm the result we obtained above.

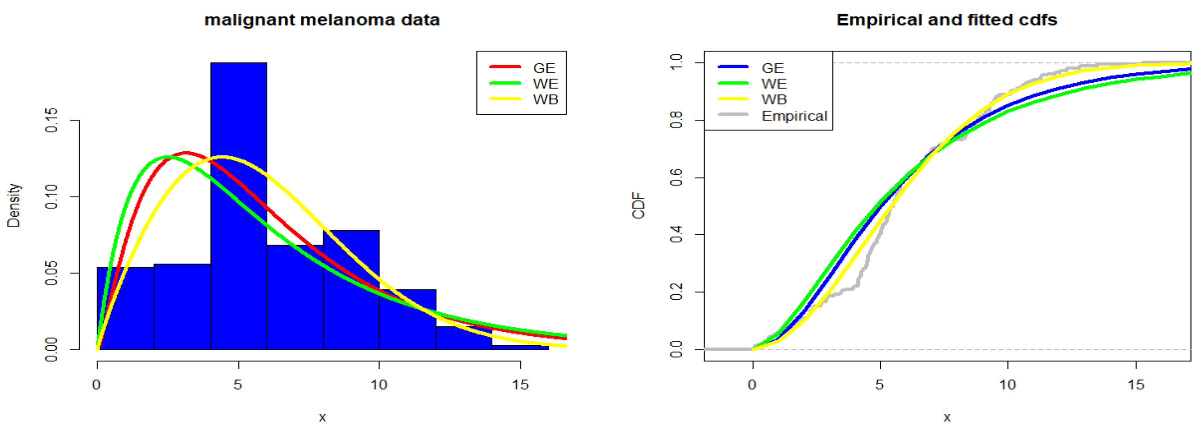

Figure 3. Histogram of the malignant melanoma data sets and the density functions of the fitted GE, WE, and WB models on the left, and estimated cdfs for the data set of the malignant melanoma on the right.

4.2. Daily Ozone Data

The second data set records the daily measurements of ozone concentration in the atmosphere of New York City from May to September 1973; for details, see Nadarajah [27]. The data are given in Appendix A.

In Table 8, we obtain the results by applying the RML, MKD, and SPRT methods. We first consider the RML criterion to analyze these real data. Since $T_n^1 = 0.4501 > 0$ and $T_n^2 = -1.6644 < 0$, we can say that the WE model fits best. Next, the MKDs between the empirical distribution function and the fitted GE, WE, and WB distribution functions are 0.0846, 0.0564 and 0.0900, respectively. Therefore, the MKD criterion also suggest choosing the WE model for this data set. The RML and the MKD criteria all lead to the same conclusion. Finally, we use the SPRT procedure with boundaries $a_j = 2.3026$, $b_j = -2.3026$, and j = 1, 2, 3. Then, $\Lambda_t^1 = -2.3088 < -2.3026$ and $\Lambda_t^2 = 2.3112 > 2.3026$, which mean that we still prefer the WE distribution rather than the GE and WB distributions. Moreover, the average sample size $E(t)$ is about 93, which is less than n_F=116, giving 19.83% savings in the average sample size.

Table 8. Summary of the RML, the MKD, the SPRT procedures, and the goodness of fit test of the daily ozone data.

Model	GE	WE	WB
shape parameters' estimation	$\hat{\gamma} = 1.7960$	$\hat{\eta} = 9.4201$	$\hat{\beta} = 1.3402$
scale parameters' estimation	$\hat{\theta} = 0.0337$	$\lambda = 0.0205$	$\hat{\xi} = 0.0217$
T_n^1		0.4501	
T_n^2		−1.6644	
T_n^3		−1.2143	
Λ_t^1		−2.3088	
Λ_t^2		2.3112	
Λ_t^3		2.0036	
AIC	1086.7920	1085.8920	1089.2210
BIC	1092.2990	1091.3990	1094.7280
MKD	0.0846	0.0564	0.0900

The histogram in Figure 4 indicates that a positively skewed distribution may fit the data well. We provide the histogram of data set 2 and the pdfs of the GE, WE, and WB fitted distributions in Figure 4. From Figure 4, it appears that the WE distribution provides a better fit than the GE and the WB distributions in this case. Finally, we also conduct the goodness of fit test on the daily ozone data set, and the results are shown in Table 8. It is found that the WE distribution fits best.

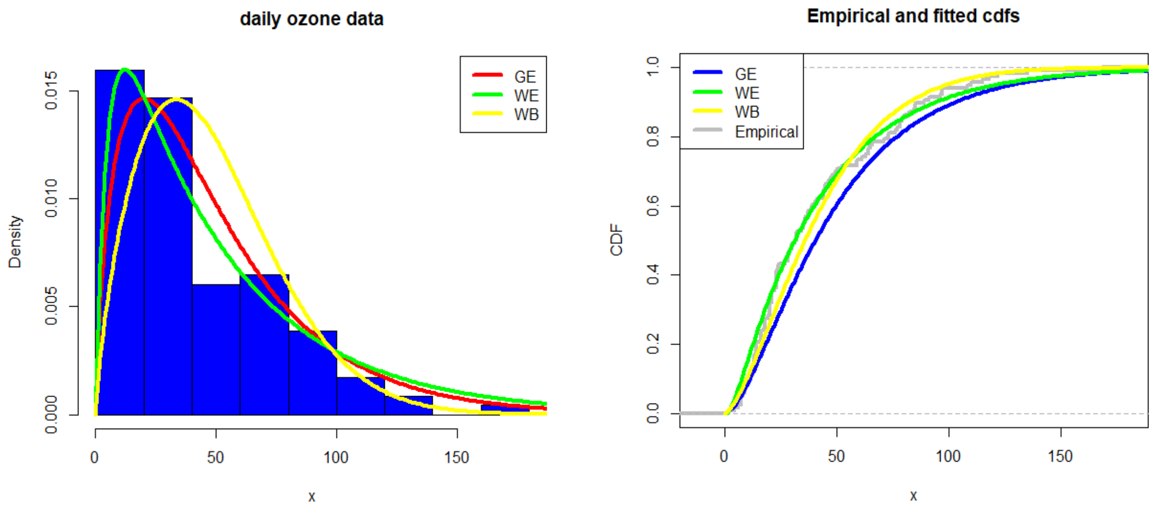

Figure 4. Histogram of the daily ozone data set and the density functions of the fitted GE, WE, and WB models on the left, and estimated cdfs for the data set of daily ozone on the right.

5. Conclusions

In this paper, we consider the problem of discriminating among the WE, the GE, and the WB distributions using three different methods. Simulation studies are conducted for the various shapes and scales of parameters, and the performance of the RML, MKD, and SPRT methods are analyzed and compared. We compute their corresponding values of PCS based on the Monte Carlo simulation for different sceneries. Through simulation studies, we find that the RML method is superior to the MKD method in distinguishing these three distributions under the fixed sample size. At the same time, without losing the values of PCS, we compare the RML method with the SPRT procedure and find that the SPRT method provides savings on the sample size. The results for the application show that the SPRT requires fewer observations to obtain the same PCS than the existing fitting test with the RML and the MKD.

Author Contributions: W.T.: conceptualization, methodology, validation, investigation, resources, supervision, project administration, visualization, and writing—review and editing; R.N.: software, formal analysis, data curation, writing—original draft preparation, and visualization; Y.Z.: software, writing—review and editing, and visualization. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Natural Science Foundation of Top Talent of SZTU under grant GDRC202214.

Data Availability Statement: The data presented in this study are openly available in [25–27].

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

Appendix A

Appendix A.1. Malignant Melanoma Data

0.0274, 0.0822, 0.0959, 0.2712, 0.5068, 0.5589, 0.5753, 0.6356, 0.6356, 0.7644, 0.8082, 0.9726, 1.0575, 1.1671, 1.2849, 1.3507, 1.4493, 1.7014, 1.7233, 1.8055, 1.8274, 1.9671, 2.0603, 2.1342, 2.1726, 2.2384, 2.2630, 2.2822, 2.3507, 2.3808, 2.3890, 2.6493, 2.6767, 2.6904, 2.8521, 2.8904, 2.9096, 2.9452, 3.1671, 3.3644, 3.4301, 3.4822, 3.5945, 3.9096, 3.9315, 4.1068, 4.1260, 4.1315, 4.1370, 4.1425, 4.1534, 4.1781, 4.2247, 4.2411, 4.2658, 4.2740, 4.2822, 4.3397, 4.3973, 4.4411, 4.4575, 4.4767, 4.4959, 4.4959, 4.5151, 4.5260, 4.5315, 4.5315, 4.5671, 4.5973, 4.6164, 4.6301, 4.6849, 4.6849, 4.7288, 4.7808, 4.8274, 4.8740, 4.8959, 4.8959, 4.9123, 4.9425, 4.9644, 5.0301, 5.0384, 5.0384, 5.0795, 5.0849, 5.0959, 5.1068, 5.2027, 5.2438, 5.2575, 5.2603, 5.2795, 5.2959, 5.3205, 5.3562, 5.3589, 5.3644, 5.3781, 5.3973, 5.4932, 5.4986, 5.5096, 5.5452, 5.5562, 5.5836, 5.6329, 5.6411, 5.6466, 5.6493, 5.6849, 5.7123, 5.7589, 5.7616, 5.7644, 5.7753, 5.7863, 5.8904, 5.9068, 5.9315, 6.0521, 6.1014, 6.1014, 6.1808, 6.2027, 6.4082, 6.4685, 6.5397, 6.5425, 6.5836, 6.6466, 6.6466, 6.6603, 6.7397, 6.7589, 6.8274, 6.8301, 6.9068, 6.9644, 7.0110, 7.0274, 7.0411, 7.2877, 7.3041, 7.3315, 7.5014, 7.6219, 7.6356, 8.1753, 8.3068, 8.3288, 8.3342, 8.4027, 8.4356, 8.4959, 8.6137, 8.6356, 8.6411, 8.7123, 8.7178, 8.7260, 8.7644, 8.8438, 8.8466, 8.9808, 9.0329, 9.1178, 9.1233, 9.1452, 9.2685, 9.2712, 9.2740, 9.2822, 9.3205, 9.4274, 9.4740, 9.4767, 9.4767, 9.5233, 9.6521, 10.0466, 10.1233, 10.1233, 10.3452, 10.3452, 10.4932, 10.5644, 10.6082, 10.7096, 10.8712, 10.9616, 11.2411, 11.2849, 11.2986, 11.5260, 11.8082, 12.0274, 12.2712, 12.3068, 12.7890, 12.7890, 13.4959, 15.2466.

Appendix A.2. Daily Ozone Data

41, 36, 12, 18, 28, 23, 19, 8, 7, 16, 11, 14, 18, 14, 34, 6, 30, 11, 1, 11, 4, 32, 23, 45, 115, 37, 29, 71, 39, 23, 21, 37, 20, 12, 13, 135, 49, 32, 64, 40, 77, 97, 97, 85, 10, 27, 7, 48, 35, 61, 79, 63, 16, 80, 108, 20, 52, 82, 50, 64, 59, 39, 9, 16, 78, 35, 66, 122, 89, 110, 44, 28, 65, 22, 59, 23, 31, 44, 21, 9, 45, 168, 73, 76, 118, 84, 85, 96, 78, 73, 91, 47, 32, 20, 23, 21, 24, 44, 21, 28, 9, 13, 46, 18, 13, 24, 16, 13, 23, 36, 7, 14, 30, 14 ,18, 20.

Appendix B

```
# the RML procedure
for (i in 1:p)
{   logWE<-function(theta){
        beta=theta[1]
        lambda=theta[2]
  h=n*log(beta+1)-n*log(beta)+n*log(lambda)-lambda*sum(x)+sum(log(1-exp(-beta*lambda*x)))
  return(-h)   }
 result1=optim(par=c(1,1),fn=logWE,control=list(maxit=1000),hessian = FALSE)
 logGE <- function(theta){
       gamma=theta[1]
       alpha=theta[2]
 h=n*log(gamma)+n*log(alpha)+(gamma-1)*sum(log(1-exp(-alpha*x)))-alpha*sum(x)
 return(h)  }
 result2=maxLik(logLik=logGE, star=c(1,1),method=''BFGS'')
```

```
  logWB<-function(theta){
       eta=theta[1]
       delta=theta[2]
  h=n*log(eta)+n*eta*log(delta)+(eta-1)*sum(log(x))-sum((delta*x)$^$eta)
  return(h)  }
  result3=maxLik(logLik = logWB, start =c(1,1),method=‘‘BFGS’’)
  LWE=n*log(result1$par[1]+1)-n*log(result1$par[1])+n*log(result1$par[2])-result1$par[2]*sum(x)+
      sum(log(1-exp(-result1$par[1]*result1$par[2]*x)))
  LGE=n*log(result2$estimate[1])+n*log(result2$estimate[2])+(result2$estimate[1]-1)*sum(log(1-exp
      (-result2$estimate[2]*x)))-result2$estimate[2]*sum(x)
  LWB=n*log(result3$estimate[1])+n*result3$estimate[1]*log(result3$estimate[2])+(result3$estimate
     [1]-1)*sum(log(x))-sum((result3$estimate[2]*x)^(result3$estimate[1])
  if((LWE-LGE)>0 & (LWE-LWB)>0){
       li=1  }
  else{       li=0  }
  qff=rbind(qff,li)}
```

the MKD procedure

```
for(i in 1:p)
{  logWE<-function(theta){
        beta=theta[1]
        lambda=theta[2]
   h=n*log(beta+1)-n*log(beta)+n*log(lambda)-lambda*sum(x)+sum(log(1-exp(-beta*lambda*x)))
   return(h)   }
  result1=maxLik(logLik = logWE, start =c(1,1),method=‘‘SANN’’)
  x1=sort(x)
  FWE=1+(1/result1$estimate[1])*exp(-result1$estimate[2]*(result1$estimate[1]+1)*x1)
      -((result1$estimate[1]+1)/result1$estimate[1])*exp(-result1$estimate[2]*x1)
  fn=ecdf(x1)
  WEKD=max(abs(FWE-fn(x1)))
  logGE<-function(theta){
       gamma=theta[1]
       alpha=theta[2]
  h=n*log(gamma)+n*log(alpha)+(gamma-1)*sum(log(1-exp(-alpha*x)))-alpha*sum(x)
  return(h) }
  result2=maxLik(logLik = logGE, start =c(1,1),method=‘‘BFGS’’)
  FGE=(1-exp(-result2$estimate[2]*x1))^(result2$estimate[1])
  GEKD=max(abs(FGE-fn(x1)))
  logWB<-function(theta){
      eta=theta[1]
      delta=theta[2]
  h3=n*log(eta)+n*eta*log(delta)+(eta-1)*sum(log(x))-sum((delta*x)^eta)
  return(h3)  }
  result3=maxLik(logLik = logWB, start =c(1,1),method=‘‘BFGS’’)
  FWB=1-exp(-(result3$estimate[2]*x1)^(result3$estimate[1]))
  WBKD=max(abs(FWB-fn(x1)))
  if((GEKD-WEKD)>0 & (WBKD-WEKD)>0){  li=1}
  else{  li=0  }
  qff=rbind(qff,li)}
```

the SPRT procedure

```
for (i in 1:p)
{ while (b1<=S1 & S1<=a1 & b1<=S2 & S2<=a1 & b1<=S3 & S3<=a1 & k<=n){
      x <- x1[1:k]
      logWE<-function(theta){
           beta=theta[1]
```

```
          lambda=theta[2]
     h=k*log(beta+1)-k*log(beta)+k*log(lambda)-lambda*sum(x)+sum(log(1-exp(-beta*lambda*x)))
     return(-h)}
result1=optim(par=c(1,1),fn=logWE,control = list(maxit=1000),hessian = FALSE)
logGE <- function(theta){
           gamma=theta[1]
           alpha=theta[2]
     h=k*log(gamma)+k*log(alpha)+(gamma-1)*sum(log(1-exp(-alpha*x)))-alpha*sum(x)
     return(h)      }
     result2=maxLik(logLik = logGE, start =c(1,1),method=''SANN'')
     logWB<-function(theta){
          eta=theta[1]
        delta=theta[2]
     h=k*log(eta)+k*eta*log(delta)+(eta-1)*sum(log(x))-sum((delta*x)^eta)
     return(h)   }
     result3=maxLik(logLik = logWB, start =c(1,1),method=''SANN'')
        s1<-function(x)
     {  P<-log(((result1$par[1]+1)/result1$par[1])*result1$par[2]*exp(-result1$par[2]*x)*
          (1-exp(-result1$par[1]*result1$par[2]*x)))
       return(P) }
     s2<-function(x)
     {    K<-log(result2$estimate[1]*result2$estimate[2]*((1-exp(-result2$estimate[2]*x))^
          (result2$estimate[1]-1))*exp(-result2$estimate[2]*x))
      return(K)  }
    s3<-function(x)
     { Q<-log(result3$estimate[1]*(result3$estimate[2]^(result3$estimate[1]))*(x^
         (result3$estimate[1]-1))*exp(-(result3$estimate[2]*x)^result3$estimate[1]))
     return(Q) }
    P=s1(x)
    K=s2(x)
    Q=s3(x)
    S1=sum(K-P)
    S2=sum(P-Q)
    S3=sum(K-Q)
    k<-k+1 }
  if(S1>=a1 & S3>=a1){       li=1   }
  else{       li=0 }
  qff=rbind(qff,li)
  N=rbind(N,k)}
```

References

1. Atkinson, A.C. A test for discriminating between models. *Biometrika* **1969**, *56*, 337–347. [CrossRef]
2. Atkinson, A.C. A method for discriminating between models. *J. R. Stat. Soc. Ser. B Methodol.* **1970**, *32*, 323–345. [CrossRef]
3. Chen, W. On the tests of separate families of hypotheses with small sample size. *J. Stat. Comput. Simul.* **1980**, *11*, 183–187. [CrossRef]
4. Chambers, E.A.; Cox, D.R. Discrimination between alternative binary response models. *Biometrika* **1967**, *54*, 573–578. [CrossRef] [PubMed]
5. Cox, D.R. Tests of separate families of hypotheses. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1960; Statistical Laboratory of the University of California: Berkeley, CA, USA, 1961; Volume 1, pp. 105–123.
6. Cox, D.R. Further results on tests of separate families of hypotheses. *J. R. Stat. Soc. Ser. B Methodol.* **1962**, *24*, 406–424. [CrossRef]
7. Dyer, A.R. Discrimination procedures for separate families of hypotheses. *J. Am. Stat. Assoc.* **1973**, *68*, 970–974. [CrossRef]
8. Weibull, W. A statistical theory of the strength of materials. *R. Swed. Inst. Eng. Res.* **1939**, *151*, 7.
9. Gupta, R.D.; Kundu, D. Theory & methods: Generalized exponential distributions. *Aust. N. Z. J. Stat.* **1999**, *41*, 173–188.
10. Gupta, R.D.; Kundu, D. A new class of weighted exponential distributions. *Statistics* **2009**, *43*, 621–634. [CrossRef]
11. Makhdoom, I. Estimation of r = p[y<x] for weighted exponential distribution. *J. Appl. Sci.* **2012**, *12*, 1384–1389.

12. Dey, S.; Ali, S.; Park, C. Weighted exponential distribution: Properties and different methods of estimation. *J. Stat. Comput. Simul.* **2015**, *85*, 3641–3661. [CrossRef]
13. Tian, W.; Yang, Y. Change point analysis for weighted exponential distribution. *Commun. Stat.-Simul. Comput.* **2021**, 1–13.. [CrossRef]
14. Gupta, R.D.; Kundu, D. Closeness of gamma and generalized exponential distribution. *Commun. Stat.-Theory Methods* **2003**, *32*, 705–721. [CrossRef]
15. Gupta, R.D.; Kundu, D. Discriminating between Weibull and generalized exponential distributions. *Comput. Stat. Data Anal.* **2003**, *43*, 179–196. [CrossRef]
16. Pakyari, R. Discriminating between generalized exponential, geometric extreme exponential and Weibull distributions. *J. Stat. Comput. Simul.* **2010**, *80*, 1403–1412. [CrossRef]
17. Raqab, M.Z. Discriminating between the generalized Rayleigh and Weibull distributions. *J. Appl. Stat.* **2013**, *40*, 1480–1493. [CrossRef]
18. Elsherpieny, E.A.; Muhammed, H.Z.; Radwan, N.U.M.M. On discriminating between gamma and log-logistic distributions in case of progressive type II censoring. *Pak. J. Stat. Oper. Res.* **2017**, *13*, 157–183. [CrossRef]
19. Ahmad, M.A.; Raqab, M.Z.; Kundu, D. Discriminating between the generalized Rayleigh and Weibull distributions: Some comparative studies. *Commun. Stat.-Simul. Comput.* **2017**, *46*, 4880–4895. [CrossRef]
20. Raqab, M.Z.; Al-Awadhi, S.A.; Kundu, D. Discriminating among Weibull, log-normal, and log-logistic distributions. *Commun. Stat.-Simul. Comput.* **2018**, *47*, 1397–1419. [CrossRef]
21. Vaidyanathan, V.S.; Varghese, A.S. Discriminating Between Exponential and Lindley Distributions. *J. Stat. Theory Appl.* **2019**, *18*, 295–302. [CrossRef]
22. Paul, B.; De, S.K.; Kundu, D. A sequential sampling approach for discriminating log-normal, Weibull, and log-logistic distributions. *Commun. Stat.-Simul. Comput.* **2021**, 1–23. [CrossRef]
23. Wald, A. Sequential tests of statistical hypotheses. *Annu. Math. Stat.* **1945**, *16*, 117–186. [CrossRef]
24. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.
25. Andersen, P.K.; Borgan, O.; Gill, R.D.; Keiding, N. *Statistical Models Based on Counting Processes*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
26. Scheike, T.H.; Zhang, M.J. Analyzing competing risk data using the R timereg package. *J. Stat. Softw.* **2011**, *38*, i02. [CrossRef] [PubMed]
27. Nadarajah, S. A truncated inverted beta distribution with application to air pollution data. *Stoch. Environ. Res. Risk Assess.* **2008**, *22*, 285–289. [CrossRef]