On maximum likelihood estimates of a proportional hazard rate model parameters based on record values

Zoran VIDOVIĆ[1]

University of Belgrade, Teacher Education Faculty, Kraljice Natalije 43,

Belgrade, Serbia

Abstract

The proportional hazard rate models have wide acceptance in modeling complex data from different engineering, reliability and survival applications. The hazard rate function of such models is able to model data with a monotonic or a bathtub-shape hazard rate. In this paper, we propose a subclass of proportional hazard rate models with three-parameters for which we have proved the existence and uniqueness of the maximum likelihood estimators based on upper kth record values. The proposed inference is found especially useful in some real illustrations within mechanical and medical analysis. Finally, several remarks are presented in the final stage of this paper.

keywords: Proportional hazard rate model, records, Hessian matrix, uniqueness, existence.
MSC(2010): 62F10, 62H12, 62F30

1 Introduction

Records are considered as those observations that overpass all of their sort. Their popularity shows stable and positive trend over the last several decades especially when events related to sports, climatology or finance are considered. Record values were first mentioned in [1] under the assumption of independent and identically distributed(iid) observations, with underlying continuous distribution. We could say that the ith observation from a sequence of iid observations is an upper record value if it exceeds all previous observations in a sequence, i.e., if $X_i > X_j$ for all $i > j$. Lower records are defined similarly. Trivially, the first observation is the first upper and lower record.

For a fixed $k \geq 1$, let us define the sequence $U_k(n)$, $n \geq 1$, of kth upper record times of iid random variables $\{X_n,\, n \geq 1\}$ with continuous distribution function $F(x)$ as follows:

$$\begin{aligned} U_k(1) &= 1, \\ U_k(n+1) &= \min\{j > U_k(n) \,:\, X_{k:U_k(n)+k-1} < X_{k:j+k-1}\},\, n \geq 1, \end{aligned}$$

where we denoted the jth order statistic of a sample $(X_1, X_2, \ldots, X_n)$ by $X_{j:n}$. The sequence $\{R_n,\, n \geq 1\} = \{X_{U_k(n):U_k(n)+k-1},\, n \geq 1\}$ is called the sequence of kth upper

[1]zoran.vidovic@uf.bg.ac.rs

record values of $\{X_n,\, n \geq 1\}$. In [2], these records are denoted as Type 2 kth record sequence, but they were first introduced in [3]. Clearly, the ordinary upper record times and values correspond to the special case when $k = 1$. For $n \geq 1$, the joint density function of the first n kth upper records is obtained as [2]

$$f_{1,\ldots,n}(r_{1(k)},\ldots,r_{n(k)}) = k^n \left(\overline{F}(r_{n(k)}))\right)^k \prod_{i=1}^{n} \frac{f(r_{i(k)})}{\overline{F}(r_{i(k)})}, \tag{1}$$

for $r_{1(k)} < \ldots < r_{n(k)}$, where $\mathbf{r}_{(\mathbf{k})} = (r_{1(k)},\ldots,r_{n(k)})$ is denoted as an observed realization of $\mathbf{R}_{(\mathbf{k})} = (R_{1(k)},\ldots,R_{n(k)})$. For a comprehensive review of record value theory we refer to [2, 4].

Let X be a random variable with support on $(0,\infty)$ from the family of continuous distribution functions with cumulative distribution function(cdf) and probability density function(pdf)

$$F(x;a,\beta,\lambda) = 1 - (\overline{G}(x;\,\beta,\lambda))^a, \quad x > 0, a > 0, \beta > 0, \lambda > 0, \tag{2}$$

and

$$f(x;a,\beta,\lambda) = ag(x;\beta,\lambda)(\overline{G}(x;\,\beta,\lambda))^{a-1}, \quad x > 0, a > 0, \beta > 0, \lambda > 0, \tag{3}$$

respectively, where $\overline{G}(x;\,\beta,\lambda) = 1-G(x;\beta,\lambda)$ and where $G(x;\beta,\lambda)$ is the cdf of a baseline random variable and a as the shape parameter. Further, let us consider the case when the baseline distribution G belongs to the class $\mathcal{C}$ of all two-parameter distributions $G(\cdot;\beta,\lambda)$ with scale parameters $\beta > 0$ and $\lambda > 0$ such that the function

$$Q(x,y,\beta,\lambda) = \frac{g(x;\beta,\lambda)}{\overline{G}(x;\beta,\lambda)(-\ln\overline{G}(y;\beta,\lambda))} I\{0 < x \leq y\} \tag{4}$$

is concave with respect to each parameter β and λ, decreases with respect to x and goes to zero as $\beta(\text{or}\,\lambda)$ goes to infinity when x and $\lambda(\text{or}\,\beta)$ are considered to be fixed.

Distribution model (2) is a special case of the proportional hazard rate (PHR) model and includes several three-parameter lifetime distributions such as modified Weibull distribution [5] and extended Weibull distribution [6]. These models are natural and tractable generalizations of the Weibull model which extends the Weibull two-parameter model by incorporating additional parameter to generate more flexible distributions, see e.g. [7, 8], and provide a better choice, having the principle of parsimony in mind, than mixtures of distributions, see e.g. [9, 10].

Various techniques in dealing with parameter estimation issues are being proposed as new extensions of distribution models occur. Certainly, the maximum likelihood procedure is the most popular providing us with maximum likelihood estimators(MLEs). In this case, scientist may deal with several problems that rise underneath this procedure. If the solution for the MLEs exist, but we cannot obtain them in explicit forms, we must rely on iteration procedures. However, iteration results may vary based on different starting values. Also, the issue of multiple roots that correspond to multiple relative maxima of the likelihood function may prove to be a large obstacle. We refer the readers for more insight on this issues to [11, 12]. In this paper, for a PHR model (3), we overcome this problem by providing sufficient conditions to check when a wide class of distributions have unique parameter MLEs based on records. We then proceeded to use such theorem on a subfamily of distributions which belongs to the model (3) on two real data sets.

Destructive stress-testing often produces efficient sampling schemes such as record series. Examples could be found in various mechanical, thermal, medical and reliability analysis where sampling of an experiment is conducted sequentially, e.q. [13]. Therefore, intuitively, the information found in a record sample is considered to be approximately equal to the information found in iid samples, [14]. For such reasons, it is important that one has developed inference based on record values. Moreover, as was noticed in [15], it may be reasonable to expect that the PRH model (3) adequately fits the data obtained from such investigations due to its wide fitting abilities.

There is an ongoing interest in the study of MLEs of PHR models. MLEs of the modified Weibull distribution and extended Weibull distribution, as special cases of PHR model (2), were discussed as well with some inferential methods for these families of distributions in [5] and [6], respectively, based on censored and complete samples. There is much research on the modified Weibull distribution based on Type-II censored samples. For example, it was proven in [16] the existence and uniqueness of the MLEs of modified Weibull distribution based on Type-II censored samples. Also, [17] discussed the performances of the least-square and maximum likelihood estimation of the parameters of modified Weibull distribution based on the same sample scheme, with respect to bias and mean square error. [18] showed that the MLEs of modified Weibull distribution uniquely exist for kth record samples. The procedure underneath considers the behaviour of the likelihood function based on records and methods from [11]. Motivated by this, the primary objective of this paper is to investigate the existence and uniqueness of the MLEs of PHR models that belong to class $\mathcal{C}$ based on kth records. Therefore, results presented in this paper can be recognized as a significant improvement over the recently published results on this topic.

The paper is organized as follows. Section 1 introduces the origin and the concept of the leading problem. Section 2 describes the necessary conditions for proving the unique existence of the MLEs of the PHR model (2) based on record data and derives the main results. Then, in Section 3 we apply these results on the analysis of two real and one simulated data examples. Section 4 concludes this article. In Section 5 we give the proofs of the main results.

2 Behavior of the likelihood function

Let $\mathbf{r}_{(\mathbf{k})} = (r_{1(k)}, r_{2(k)}, ..., r_{n(k)})$ be a sample of kth records following (2), $n \geq 2$ and fixed $k \geq 1$. Their joint distribution and its logarithm are, respectively, given as

$$L(a, \beta, \lambda; \mathbf{r}_{(\mathbf{k})}) = k^n a^n \left(\overline{G}(x; \beta, \lambda)\right)^k \prod_{i=1}^{n} \frac{g(r_{i(k)}; \beta, \lambda)}{\overline{G}(r_{i(k)}; \beta, \lambda)}, \tag{5}$$

$$\ln L(a,\beta,\lambda;\mathbf{r_{(k)}}) = n\ln a + n\ln k + ka\ln(\overline{G}(r_{n(k)};\beta,\lambda)) + \sum_{i=1}^{n}\ln g(r_{i(k)};\beta,\lambda)$$
$$-\sum_{i=1}^{n}\ln\overline{G}(r_{i(k)};\beta,\lambda). \tag{6}$$

The MLE of α is obtained from the likelihood equation by using partial derivative with respect to α and equating it to zero, i.e.,

$$\frac{\partial \ln L}{\partial a} = \frac{n}{a} + k\ln(\overline{G}(r_{n(k)};\beta,\lambda)) = 0.$$

It is convenient to denote the function

$$a(\beta,\lambda) = \frac{n}{-k\ln(\overline{G}(r_{n(k)};\beta,\lambda))}, \tag{7}$$

for $\beta > 0$ and $\lambda > 0$.

Then, we have that the MLE for a, when MLEs $\hat{\beta}$ and $\hat{\lambda}$ are known, could be expressed as

$$\hat{a} = a(\widehat{\beta},\widehat{\lambda}) = \frac{n}{-k\ln(\overline{G}(r_{n(k)};\widehat{\beta},\widehat{\lambda}))}. \tag{8}$$

Based on the inequality $\ln t \le t-1$, $t>0$, [19] assures us that the inequality

$$\ln L(a,\beta,\lambda;\mathbf{r_{(k)}}) \le L^*(\beta,\lambda), \text{ for all } \beta,\ \lambda > 0, \tag{9}$$

holds, with equality for $a = a(\beta,\lambda)$, where

$$L^*(\beta,\lambda) = n\ln n - n - n\ln(-\ln(\overline{G}(r_{n(k)};\beta,\lambda))) + \sum_{i=1}^{n}\ln g(r_{i(k)};\beta,\lambda)$$
$$-\sum_{i=1}^{n}\ln\overline{G}(r_{i(k)};\beta,\lambda). \tag{10}$$

It can be seen straightforwardly that by maximizing the function $L^*(\beta,\lambda)$ we directly maximize the loglikelihood function $L(a,\beta,\lambda;\mathbf{r_{(k)}})$, as was also nicely deduced in [16, 18] with a similar problem at hand. Moreover, it may be realized that function L^* does not depend on the parameter k.

Next, for further progress we denote the extended parameter space as $\Theta_0^* = \{(a,\beta,\lambda) : a > 0,\ g(r_{i(k)};\beta,\lambda) > 0,\ i = 1,...,n\}$ of the parameter space $\Theta_0 = \{(a,\beta,\lambda) : a > 0, \beta > 0, \lambda > 0\}$. Also, let us denote $\Theta^* = \{(\beta,\lambda) : g(r_{i(k)};\beta,\lambda) > 0,\ i = 1,...,n\}$. It is quite evident that Θ_0^* and Θ^* have a common boundary, i.e., $\partial\Theta_0^* = \partial\Theta^* = \{\beta = \infty\} \cup \{\lambda = \infty\} \cup \Theta_1^* \cup \Theta_2^*$, where $\Theta_1^* = \{(\beta,\lambda) : \beta > 0,\ g(r_{i(k)};\beta,\lambda) > 0,\ i = 1,...,n\}$ and $\Theta_2^* = \{(\beta,\lambda) : \lambda > 0,\ g(r_{i(k)};\beta,\lambda) > 0,\ i = 1,...,n\}$. This will come in handy in what follows.

The approach from [11] requires that the following two conditions

(1) $\lim_{(\beta,\lambda)\to\partial\Theta^*} L^*(\beta,\lambda) = -\infty$,

(2) the Hessian matrix $H(\beta,\lambda)$ is negative definite at every point $(\beta,\lambda)\in\Theta^*$,

must be satisfied in order to ensure the existence and uniqueness of the MLEs of (β,λ). For this purpose, next technical lemmas will be useful.

Lemma 2.1 *For fixed $k\geq 1$ and $n\geq 2$,*

$$\lim_{\lambda\to\infty}\sup_{(\beta,\lambda)\in\Theta_1^*} L^*(\beta,\lambda)=-\infty.$$

Proof. The proof is given in Appendix A.

Lemma 2.2 *For fixed $k\geq 1$,*

$$\lim_{\beta\to\infty}\sup_{(\beta,\lambda)\in\Theta_2^*} L^*(\beta,\lambda)=-\infty.$$

Proof. The proof follows similar steps as in Lemma 2.1.

Lemma 2.3 *For fixed $k\geq 1$ and $n\geq 2$,*

$$\lim_{\lambda\to\infty}\sup_{\beta>0:\,g(r_{i(k)};\beta,\lambda)=0,\,for\ some\,i=\overline{1,n}} L^*(\beta,\lambda)=-\infty.$$

Proof. The proof follows directly from the behaviour of the logarithmic function.

Lemma 2.4 *For fixed $k\geq 1$ and $n\geq 2$,*

$$\lim_{\beta\to\infty}\sup_{\lambda>0:\,g(r_{i(k)};\beta,\lambda)=0,\,for\ some\,i=\overline{1,n}} L^*(\beta,\lambda)=-\infty.$$

Proof. The proof is similar to Lemma 2.3.

The next theorem is obtained directly from Lemmas 2.2-2.4.

Theorem 2.5 *If $\lim_{(\beta,\lambda)\to\partial\Theta^*} L^*(\beta,\lambda)=-\infty$, the log-likelihood function is constant on the boundary $\partial\Theta^*$ of the parameter space Θ^*.*

In order to prove the following lemma, we must assume that the function $L^*(\beta,\lambda)$ satisfies the relation

$$\frac{\partial^2 L^*}{\partial\beta^2}\frac{\partial^2 L^*}{\partial\lambda^2}>\left(\frac{\partial^2 L^*}{\partial\beta\lambda}\right)^2. \tag{11}$$

This condition could be proven via a semi-parametric approach see e.g.[20].

Lemma 2.6 *The Hessian matrix $H(\beta,\lambda)$ is negative-definite for all $(\beta,\lambda)\in\Theta^*$.*

Proof. Consider the properties of class $\mathcal{C}$ and (11). See [18] for more details.

We can now state the final theorem of this paper.

Theorem 2.7 *When considering upper kth record values, the MLEs of a,β and λ of a population with (2) exists and are unique in the parameter space Θ_0^*.*

This theorem confirms the unique existence of the MLEs on the parameter space Θ_0^*. On the other hand, some technical complexities during this process may occur. One of the main concerns is the case when the numerically estimated MLEs fall out of the first quarter of $\beta-\lambda$ space. The intuitive solution for such issue is to consider it as an constrained optimization problem on Θ^*, as was done in [16]. The MLEs obtained in such way preserve all regularity conditions and, in the same time, maximize the loglikelihood function.

3 Examples

Here, we provide two examples to illustrate all methods of inference discussed before. First example was nicely deduced in [18] and deals with unique estimation of MLEs for the modified Weibull distribution based on records, which belongs to class $\mathcal{C}$. Another example of a PHR model that belongs to the class $\mathcal{C}$ and satisfies (11) has a cumulative distribution function(cdf), probability density function(pdf) and hazard rate function(hrf) as

$$F(x;\alpha,\beta,\lambda) = 1 - e^{-\alpha x^{\beta} e^{-\frac{\lambda}{x}}}, \tag{12}$$

$$f(x;\alpha,\beta,\lambda) = \alpha(\lambda+\beta x)x^{\beta-2}e^{-\frac{\lambda}{x}-\alpha x^{\beta}e^{-\frac{\lambda}{x}}}, \tag{13}$$

$$h(x;\alpha,\beta,\lambda) = \alpha(\lambda+\beta x)x^{\beta-2}e^{-\frac{\lambda}{x}}, \tag{14}$$

for $x > 0, \alpha > 0, \beta > 0, \lambda > 0$.

This distribution model is known as the three-parameter extended Weibull model which was proposed in [6], which we shall denote as $PYEW(a,\beta,\lambda)$. It is recognized that when $0 < \beta < 1$, this distribution suits to model lifetime data with upside-down bathtub shape hazard rate, while for $\beta \geq 1$ it has the potential to model data with increasing hazard rates. Its ability to model data sets with upside-down bathtub shape hazard function, with a minimum parameters used, makes it an attractive model for researchers. Therefore, this distribution could be recognized as a potential model for various reliability testing, human life or mechanical devices lifetime data.

Given observed kth upper record values $\mathbf{r_{(k)}} = (r_{1(k)}, r_{2(k)}, ..., r_{n(k)})$, for this case (5) and (6) reduce to:

$$L(\alpha,\beta,\lambda;\mathbf{r_{(k)}}) = \alpha^n k^n h(\alpha,\beta,\lambda;\mathbf{r_{(k)}})e^{-\alpha T(\beta,\lambda;\mathbf{r_{(k)}})}, \tag{15}$$

and

$$\ln L(\alpha,\beta,\lambda;\mathbf{r_{(k)}}) = n\ln\alpha + n\ln k + \ln h(\alpha,\beta,\lambda;\mathbf{r_{(k)}}) - \alpha T(\beta,\lambda;\mathbf{r_{(k)}}), \tag{16}$$

where

$$h(\alpha,\beta,\lambda;\mathbf{r_{(k)}}) = \prod_{i=1}^{n}(\lambda+\beta r_{i(k)})r_{i(k)}^{\beta-2}e^{-\frac{\lambda}{r_{i(k)}}}$$

and

$$T(\beta,\lambda;\mathbf{r_{(k)}}) = k r_{n(k)}^{\beta} e^{-\frac{\lambda}{r_{n(k)}}}.$$

The MLE's of the parameters α,β and λ, denoted as $\widehat{\alpha},\widehat{\beta}$ and $\widehat{\lambda}$, respectively, are obtained from the system of likelihood equations by using partial derivatives with respect to each parameter and equating them to zero, i.e.,

$$\begin{aligned}
\frac{\partial \ln L}{\partial \alpha} &= \frac{m}{\alpha} - T(\beta,\lambda;\mathbf{r_{(k)}}) = 0,\\
\frac{\partial \ln L}{\partial \beta} &= -\alpha r_{n(k)}^{\beta} \ln r_{n(k)} e^{-\frac{\lambda}{r_{n(k)}}} + \sum_{i=1}^{n} \frac{r_{i(k)}}{(\lambda+\beta r_{i(k)})} + \sum_{i=1}^{n} \ln r_{i(k)} = 0, \qquad (17)\\
\frac{\partial \ln L}{\partial \lambda} &= -\sum_{i=1}^{n} \frac{1}{r_{i(k)}} + \sum_{i=1}^{n} \frac{1}{(\lambda+\beta r_{i(k)})} + \alpha k r_{n(k)}^{\beta-1} e^{-\frac{\lambda}{r_{n(k)}}} = 0.
\end{aligned}$$

Then, we have that the MLE for α, when MLEs $\hat{\beta}$ and $\hat{\lambda}$ are known, can be expressed as

$$\hat{\alpha} = a(\widehat{\beta},\widehat{\lambda}) = \frac{n}{T(\widehat{\beta},\widehat{\lambda};\mathbf{r_{(k)}})}. \qquad (18)$$

Also, for this case, the function (10) reduces to

$$\begin{aligned}
L^{*}(\beta,\lambda) = n\ln n - n - n\beta \log r_{n(k)} + \frac{n\lambda}{r_{n(k)}} + \sum_{i=1}^{n} \ln(\lambda+\beta r_{i(k)})\\
+ (\beta-2)\sum_{i=1}^{n} \ln r_{i(k)} - \lambda \sum_{i=1}^{n} \frac{1}{r_{i(k)}}. \qquad (19)
\end{aligned}$$

Based on (19), it is evident that Theorem 2.7 holds for this case.

3.1 Illustrations

In this part of the paper, we analyze the uniqueness of the MLEs based on records from practical data sets with the fitting model (12) and using a simulated data set from the same model with given parameters. All numerical analysis were done in a open source statistical program R [21], using four digit accuracy.

3.1.1 Data set 1.

The first real data set is presented in Table 1 and found in [22, pg. 574] and [23]. It gives us the survival times of steel speciments undergoing stress level 32.0 under mechanical testing. We analyzed the complete data by displaying the Weibull Probability Plot (WPP) and Total Time on Test (TTT) plots based on it. These plots are presented in Figure 1.

Table 1: Lifetimes of steel speciment with stress level 32.0.

1144	231	523	474	4510	3107	815	6297	1580	605	1786
206	1943	935	283	1336	727	370	1056	413	619	2214
1826	597									

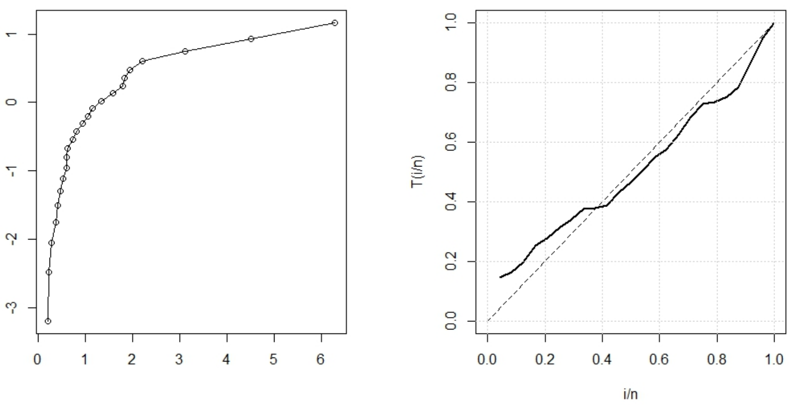

Figure 1: WPP and TTT plot.

The WPP and TTT plots show that the model (12) can be reasoned as a correct fitting distribution, what was indicated in [6]. Its validity was confirmed by a Lilliefors correction of the Kolmogorov-Smirnov (K-S) test with 10000 replications.

The MLEs of the model parameters, estimated K-S distance between the fitted and the empirical distribution function and the corresponding estimate of the p-value are computed numerically and presented in Table 2. The results indicate that model (12) fits the data set well.

Table 2: MLEs of the $PYEW$ model parameters, estimated K-S distance and p-value.

$PYEW(\hat{a}, \hat{\beta}, \hat{\lambda})$	K-S	p-value
$(0.0128, 0.6671, 498.1398)$	0.0757	0.9973

Let us first extract the kth records from the data set. They are as follows:

i	1	2	3	4	5	6	7	8	9
$R_{i(1)}$	1144	4510	6297						
$R_{i(2)}$	231	523	1144	3107	4510				
$R_{i(3)}$	231	474	523	1144	3107				
$R_{i(4)}$	231	474	523	815	1144	1580	1786	1943	2214

Let us select the values $k = 1$ and $k = 2$ for our analysis, which we shall denote as Case I and Case II, respectively.

- **Case I** Based on the optimization procedure on Θ_0, we observe that the MLEs of (β, λ) are obtained at the point $(1.2552, 565.8382)$. The joint MLEs are, therefore, $(\hat{a}, \hat{\beta}, \hat{\lambda}) = (0.0001, 1.2552, 565.8382)$. At this point the value -25.7299 is declared as the maximum of the log-likelihood function (10). For a detailed overview, it is noteworthy to illustrate the behavior of log-likelihood function (19) on the area $[0, 2] \times [100, 900]$. In Figure 2 we illustrate the appropriate behavior pattern via a surface and a contour plot. We can see that the log-likelihood function follows a quasi-tunnel shape. This confirms that the MLEs are unique.

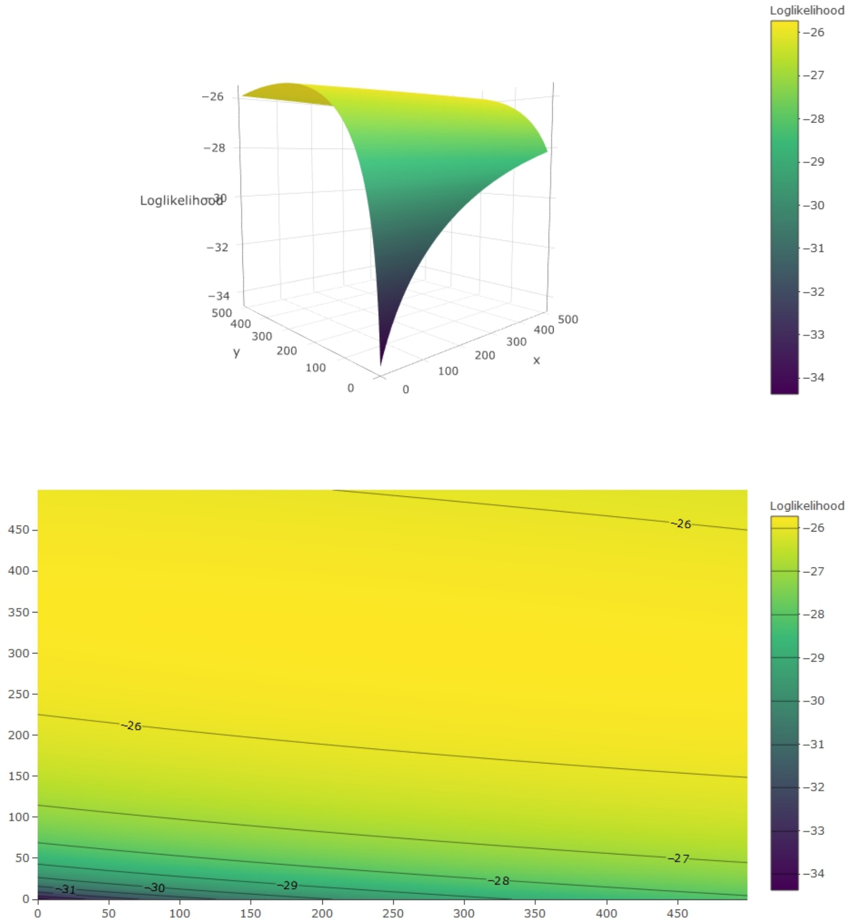

Figure 2: Surface and contour plot for Case I.

- **Case II**

  Following the same steps, we yield the MLEs $(\hat{\beta}, \hat{\lambda}) = (0.423, 319.6865)$ with -38.2288 as the associated maximum of (19). The joint MLEs for this case are $(\hat{a}, \hat{\beta}, \hat{\lambda}) = (0.0763, 0.423, 319.6865)$. Further, we can focus on the behavior of the log-likelihood function (19). According to the surface and contour plot on $[0, 1] \times [100, 500]$ illustrated in Figure 3, we can see a similar behavior pattern as in previous case. This allows us to conclude that the MLEs are unique.

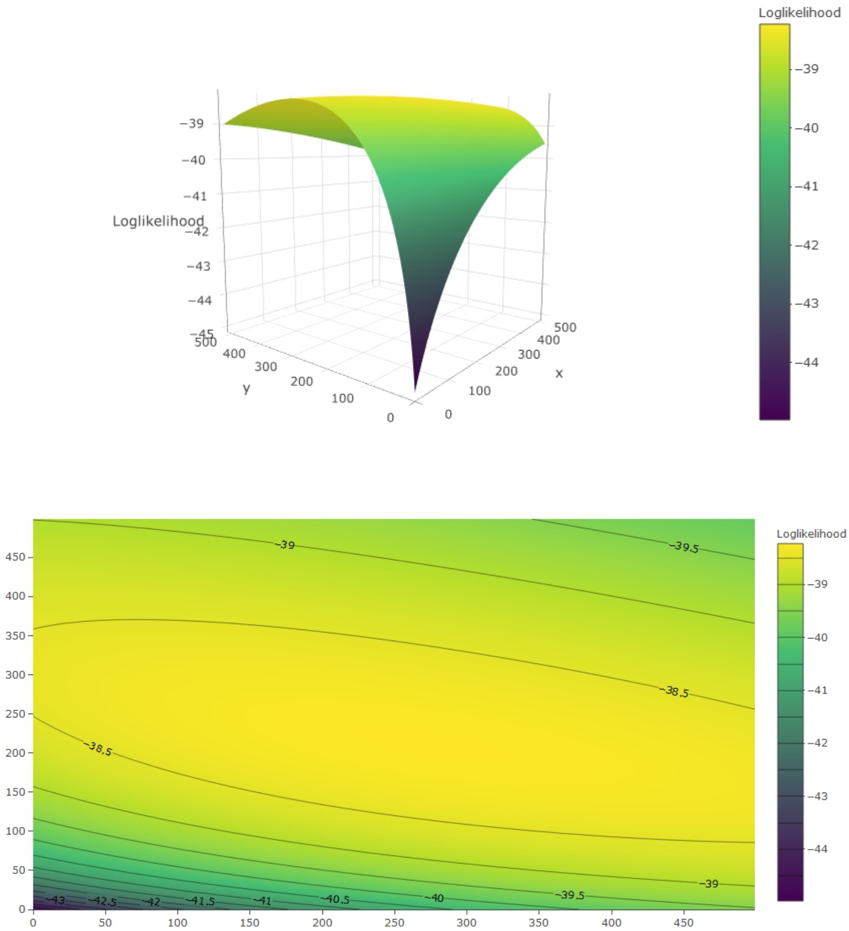

Figure 3: Surface and contour plot for Case II.

3.1.2 Data set 2.

The second data set consists on remission times (in months) of a random sample of 128 bladder cancer patients found in [24]. The data is given in Table 3.

Table 3: Remission times (in months) of 128 bladder cancer patients.

0.08	2.09	3.48	4.87	6.94	8.66	13.11	23.63
0.20	2.23	3.52	4.98	6.97	9.02	13.29	0.40
2.26	3.57	5.06	7.09	9.22	13.80	25.74	0.50
2.46	3.64	5.09	7.26	9.47	14.24	25.82	0.51
2.54	3.70	5.17	7.28	9.74	14.76	26.31	0.81
2.62	3.82	5.32	7.32	10.06	14.77	32.15	2.64
3.88	5.32	7.39	10.34	14.83	34.26	0.90	2.69
4.18	5.34	7.59	10.66	15.96	36.66	1.05	2.69
4.23	5.41	7.62	10.75	16.62	43.01	1.19	2.75
4.26	5.41	7.63	17.12	46.12	1.26	2.83	4.33
5.49	7.66	11.25	17.14	79.05	1.35	2.87	5.62
7.87	11.64	17.36	1.40	3.02	4.34	5.71	7.93
11.79	18.10	1.46	4.40	5.85	8.26	11.98	19.13
1.76	3.25	4.50	6.25	8.37	12.02	2.02	3.31
4.51	6.54	8.53	12.03	20.28	2.02	3.36	6.76
12.07	21.73	2.07	3.36	6.93	8.65	12.63	22.69

WPP and TTT plots indicate that $PYEW$ model could be selected as the appropriate fitting model whereas the information that consists on MLEs of the model parameters, estimated K-S distance and p-value, given in Table 4, confirms the model selection.

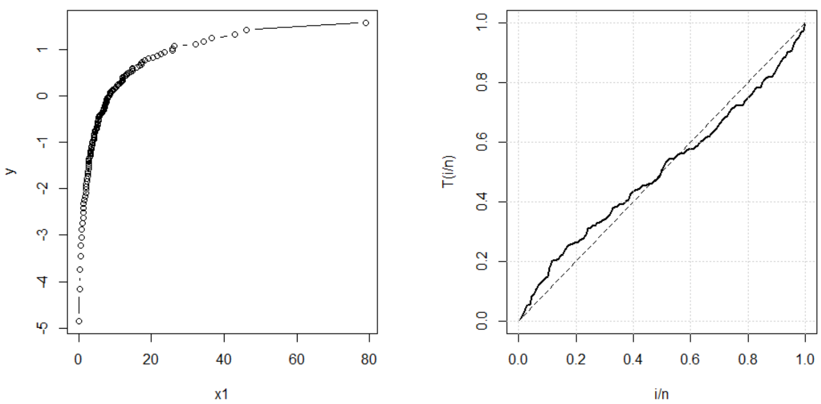

Figure 4: WPP and TTT plot.

Table 4: MLEs of the $PYEW$ model parameters, estimated K-S distance and p-value.

$PYEW(\hat{a}, \hat{\beta}, \hat{\lambda})$	K-S	p-value
$(0.1026, 1.0197, 0.1271)$	0.0697	0.5362

As above, the kth records from the data set are extracted and presented in the following:

i	1	2	3	4	5	6	7	8	9	10
$R_{i(1)}$	0.08	2.09	3.48	4.87	6.94	8.66	13.11	23.63	25.74	25.82
	26.31	32.15	34.26	36.66	43.01	46.12	79.05			
$R_{i(2)}$	0.08	2.09	3.48	4.87	6.94	8.66	13.11	13.29	13.80	23.63
	25.74	25.82	26.31	32.15	34.26	36.66	43.01	46.12		
$R_{i(3)}$	0.08	2.09	3.48	4.87	6.94	8.66	9.02	13.11	13.29	13.80
	14.24	23.63	25.74	25.82	26.31	32.15	34.26	36.66	43.01	
$R_{i(4)}$	0.08	2.09	3.48	4.87	6.94	6.97	8.66	9.02	9.22	13.11
	13.29	13.80	14.24	14.76	23.63	25.74	25.82	26.31	32.15	34.26
	36.66									

Cases $k = 1$ and $k = 2$ are of interest for our analysis. In what follows, these cases are are denoted as Case I and Case II, respectively.

- **Case I**

  By constraint optimization procedure on Θ_0 we obtain the MLEs $(\hat{\beta}, \hat{\lambda}) = (0.5247, 0.0498)$ with -39.0769 as the attained maximum of the log-likelihood function (19). The unique joint MLEs are $(\hat{a}, \hat{\beta}, \hat{\lambda}) = (1.7173, 0.5247, 0.0498)$.

- **Case II**

  Under the same procedure as above, the MLEs are $(\hat{\beta}, \hat{\lambda}) = (0.7052, 0.0312)$. At this point the value -33.7946 is declared as the maximum of the log-likelihood function (19). The unique joint MLEs are, therefore, $(\hat{a}, \hat{\beta}, \hat{\lambda}) = (0.6041, 0.7052, 0.0312)$.

3.2 Simulated data

Here, we provide a simulated random sample of size 20 from (13) with parameters $(a, \beta, \lambda) = (1, 1, 1)$ for illustration purpose. These observations are given in Table 5.

Table 5: Simulated data from $PYEW(1, 1, 1)$.

1.6140	1.4053	0.5791	1.1652	0.4257	1.2864	1.0063	2.0966	3
1.0869	4.6187	0.5521	1.3429	3.3394	2.1991	2.9478	1.6469	1.6787
1.0030	1.4239							

The kth records extracted from the simulated data set are as follows:

i	1	2	3	4	5	6	7	8
$R_{i(1)}$	1.6140	2.0966	3	4.6187				
$R_{i(2)}$	1.4053	1.6140	2.0966	3	3.3394			
$R_{i(3)}$	0.5791	1.1652	1.2864	1.4053	1.6140	2.0966	3	
$R_{i(4)}$	0.5791	1.1652	1.2864	1.4053	1.6140	2.0966	2.1991	2.9478

Let us devote our attention to cases $k = 1$ and $k = 2$ and denote them as Case I and Case II, respectively.

- **Case I**

  Constraint optimization on Θ_0 yields the MLEs $(\hat{\beta}, \hat{\lambda}) = (0, 5.1259)$ with -3.6131 as the attained maximum value of (10). However, in this setting the system (17) becomes ill-defined. So, in order to overcome this issue, we expand the parameter space Θ_0 to Θ_0^* and obtain the point $(\hat{\beta}_0, \hat{\lambda}_0) = (-0.3016, 6.0042)$ with -3.6036 as the achieved maximum value of (10). But, by this we may notice that the fitted $PYEW$ model becomes an ill-suitable model choice. Therefore, we accept the MLEs $(\hat{a}, \hat{\beta}, \hat{\lambda}) = (9.1014, 0, 5.1259)$ and recognize that the two-parameter model $F(x; a, \lambda) = 1 - e^{-ae^{-\frac{\lambda}{x}}}$ is preferable model for this case.

- **Case II**

  The same procedure holds as in previous case. The obtained MLEs $(\hat{a}, \hat{\beta}, \hat{\lambda}) = (9.7808, 0, 5.3005)$ are attached with the maximum -1.3429 of (10), and therefore we accept the two-parameter model $F(x; a, \lambda) = 1 - e^{-ae^{-\frac{\lambda}{x}}}$ as the adequate one.

4 Conclusion

In this paper, we have studied the properties of the class of three-parameter distributions for which we can confirm the existence and uniqueness of the parameter MLEs based on kth records. We proposed sufficient conditions to check if the parameter MLEs are unique for such distributions. The main theorem is used in different extensions of the two-parameter Weibull distribution. We also show the applicability of the main theorem on two real data sets and on a simulated data set. The find above is a interesting result since it could directly provide necessary information about the model parameters when record data is considered.

Moreover, it is of interest to mention here a T-X distribution family [25] with cdf

$$F(x) = \int_a^{W[H(x;\xi)]} r(t)\, dt, \tag{20}$$

where $r(t)$ is a pdf of a random variable T with support on $[a, b]$, $-\infty \leq a < b \leq \infty$, $H(x; \xi)$ is a baseline cdf with parameter vector ξ and $W[H(x; \xi)]$ is a function of the baseline cdf with properties:
(i) $W[H(x; \xi)] \in [a, b]$,
(ii) $W[H(x; \xi)]$ is differentiable and monotonically non-decreasing and
(iii) $\lim_{x\to-\infty} W[H(x; \xi)] = a$ and $\lim_{x\to\infty} W[H(x; \xi)] = b$.
The corresponding pdf is of the form

$$f(x) = \left\{ \frac{\partial}{\partial x} W[H(x; \xi)] \right\} r(W[H(x; \xi)]). \tag{21}$$

The case when T follows (12) with $W[H(x; \xi)] = \frac{H(x;\xi)}{1-H(x;\xi)}$ provides us a distribution family with cdf

$$F(x; \alpha, \beta, \lambda) = 1 - \exp\left\{ -\alpha \left(\frac{H(x; \xi)}{1 - H(x; \xi)} \right)^{\beta} e^{-\lambda\left(\frac{1-H(x;\xi)}{H(x;\xi)}\right)} \right\}, \tag{22}$$

for $x \in [a,b]$ and $\alpha > 0, \beta > 0, \lambda > 0$. The main intention when producing such models is to adapt kurtosis and skewness to a preferable level. This model was nicely discussed in [26].

In the case when cdf $H(x;\xi)$ has no unknown parameters, i.e. $H(x;\xi) = H(x)$, it can be easily shown that MLEs of parameters α, β and λ based on kth records from (22) are uniquely determined since all conditions of the class $\mathcal{C}$ are fulfilled, including (11). The model (12) is a special case of the model (22) when $H(x) = \frac{x}{x+1}$ for $x > 0$ and zero for $x \leq 0$. The same conclusions hold true when a random variable T follows a modified Weibull distribution [5], although such models are not discussed in the literature still.

5 Appendix

5.1 Proof of Lemma 2.1

For fixed β and k, we deduce that

$$L^*(\beta, \lambda) < g_\beta(\lambda),$$

where

$$g_\beta(\lambda) = n \ln n - n - n \ln(-\ln \overline{G}(r_{n(k)}; \beta, \lambda)) + \sum_{i=1}^{n} \ln g(r_{i(k)}; \beta, \lambda) - n \ln \overline{G}(r_{n(k)}; \beta, \lambda).$$

For fixed $n \geq 2$, function $g_\beta(\lambda)$ achieves its maximum at the point λ^* such that $g'_\beta(\lambda) = 0$, i.e., this implies that

$$\sum_{i=1}^{n} \frac{g(r_{i(k)}; \beta, \lambda)}{\overline{G}(r_{n(k)}; \beta, \lambda)(-\ln \overline{G}(r_{n(k)}; \beta, \lambda))}$$

is a function of β only, which we shall denote as $C(\beta)$. Therefore

$$g_\beta(\lambda^*) = n \ln n - n + C(\beta)$$

where, additionally, $C(\beta)$ can be viewed as

$$C(\beta) = \sum_{i=1}^{n} \ln \frac{\overline{G}(r_{i(k)}; \beta, \lambda)}{\overline{G}(r_{n(k)}; \beta, \lambda)} + \sum_{i=1}^{n} \ln Q(r_{i(k)}, r_{n(k)}, \beta, \lambda). \tag{23}$$

Since G belongs to $\mathcal{C}$, we have that $C(\beta) \to -\infty$ when $\lambda \to \infty$.

With this in mind, we may conclude that

$$\sup_{(\beta,\lambda)\in\Theta_1^*} L^*(\beta, \lambda) \to -\infty \text{ as } \lambda \to \infty.$$

This ends the proof.

Acknowledgments

The author would like to thank the Editor and the reviewer for careful reading and valuable comments, which led to an improved presentation of this paper.

References

[1] K. N. Chandler. The distribution and frequency of record values. *Journal of the Royal Statistical Society Series B (Methodological)*, 14, 1952.

[2] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *Records*. Wiley, 1998.

[3] W. Dziubdziela and B. Kopociński. Limiting properties of the k-th record values. *Applicationes Mathematicae*, 2(15):187–190, 1976.

[4] V. B. Nevzorov. *Records: Mathematical Theory*. Translations of Mathematical Monographs. AMS, 2000.

[5] C. D. Lai, X. Min, and D. N. P. Murthy. A modified Weibull distribution. *IEEE Transactions on reliability*, 52(1):33–37, 2003.

[6] X. Peng and Z. Yan. Estimation and application for a new extended Weibull distribution. *Reliability Engineering & System Safety*, 121, 2014.

[7] H. Pham and C. D. Lai. On recent generalizations of the Weibull distribution. *IEEE Transactions on Reliability*, 56(3):454–458, 2007.

[8] S. J. Almalki and S. Nadarajah. Modifications of the Weibull distribution: A review. *Reliability Engineering & System Safety*, 124:32–55, 2014.

[9] S. Jiang and D. Kececioglu. Graphical representation of two mixed-Weibull distributions. *IEEE Transactions on Reliability*, 41(2):241–247, 1992.

[10] H. J. Ma and Z. Z. Yan. Discrete Weibull-Rayleigh distribution properties and parameter estimations. *Thermal Science*, 26(3 Part B):2627–2636, 2022.

[11] T. Mäkeläinen, K. Schmidt, and G. P. H. Styan. On the existence and uniqueness of the maximum likelihood estimate of a vector-valued parameter in fixed-size samples. *The Annals of Statistics*, pages 758–767, 1981.

[12] V. D. Barnett. Evaluation of the maximum-likelihood estimator where the likelihood equation has multiple roots. *Biometrika*, 53(1/2):151–165, 1966.

[13] N. Glick. Breaking records and breaking boards. *American Mathematical Monthly*, pages 2–26, 1978.

[14] J. Ahmadi and N. R. Arghami. Comparing the Fisher information in record values and iid observations. *Statistics*, 37(5):435–441, 2003.

[15] Lai C. D. Bebbington, M. and R. Zitikis. Useful periods for lifetime distributions with bathtub shaped hazard rate functions. *IEEE Transactions on Reliability*, 55(2): 245–251, 2006.

[16] H. Jiang, M. Xie, and L. C. Tang. On MLEs of the parameters of a modified Weibull distribution for progressively type-2 censored samples. *Journal of Applied Statistics*, 37(4):617–627, 2010.

[17] H. K. T. Ng. Parameter estimation for a modified Weibull distribution, for progressively type-II censored samples. *IEEE Transactions on Reliability*, 54(3):374–380, 2005.

[18] Z. Vidović. On MLEs of the parameters of a modified Weibull distribution based on record values. *Journal of Applied Statistics*, 46(4):715–724, 2018.

[19] N. Balakrishnan and E. Cramer. *The art of progressive censoring*. Springer, 2014.

[20] A. Pak and S. Dey. Statistical inference for the power lindley model based on record values and inter-record times. *Journal of Computational and Applied Mathematics*, 347:156–172, 2019.

[21] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015.

[22] J. F. Lawless. *Statistical models and methods for lifetime data*, volume 362. John Wiley & Sons, 2011.

[23] M. Crowder. Tests for a family of survival models based on extremes. In *Recent Advances in Reliability Theory*, pages 307–321. Springer, 2000.

[24] E. T. Lee and J. Wang. *Statistical methods for survival data analysis*, volume 476. John Wiley & Sons, 2003.

[25] A. Alzaatreh, C. Lee, and F. Famoye. A new method for generating families of continuous distributions. *Metron*, 71(1):63–79, 2013.

[26] M. Ç. Korkmaz. A new family of the continuous distributions: The extended Weibull-G family. *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, 68(1):248–270, 2018.

Submitted: 23.08.2022.
Revised: 11.10.2022.
Accepted: 15.10.2022.