3-27-2017

Maximum Likelihood Estimation of Parameters in Exponential Power Distribution with Upper Record Values

Tianchen Zhi
Florida International University, tzhi001@fiu.edu

DOI: 10.25148/etd.FIDC001759
Follow this and additional works at: https://digitalcommons.fiu.edu/etd
Part of the Applied Statistics Commons, Other Statistics and Probability Commons, and the Statistical Models Commons

Recommended Citation
https://digitalcommons.fiu.edu/etd/3211

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.
MAXIMUM LIKELIHOOD ESTIMATION OF PARAMETERS IN EXPONENTIAL POWER DISTRIBUTION WITH UPPER RECORD VALUES

A thesis submitted in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE in STATISTICS by Tianchen Zhi

2017
To: Dean Michael R. Heithaus  
   College of Arts, Sciences and Education  

This thesis, written by Tianchen Zhi, and entitled Maximum Likelihood Estimation of Parameters in Exponential Power Distribution with Upper Record Values, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this thesis and recommend that it be approved.

_______________________________________  
Florence George

_______________________________________  
Kai Huang

_______________________________________  
Jie Mi, Major Professor

Date of Defense: March 27, 2017  

The thesis of Tianchen Zhi is approved.

_______________________________________  
Dean Michael R. Heithaus  
   College of Arts, Sciences and Education

_______________________________________  
Andrés G. Gil  
   Vice President for Research and Economic Development  
   and Dean of the University Graduate School

Florida International University, 2017
ACKNOWLEDGMENTS

I wish to thank the members of my committee for their support and patience. I want to thank my supervisor Dr. Mi for his guides through every steps of my thesis project. His broad and profound knowledge gave me great help.
ABSTRACT OF THE THESIS

MAXIMUM LIKELIHOOD ESTIMATION OF PARAMETERS IN EXPONENTIAL POWER DISTRIBUTION WITH UPPER RECORD VALUES

by

Tianchen Zhi

Florida International University, 2017

Miami, Florida

Professor Jie Mi, Major Professor

The exponential power (EP) distribution is a very important distribution that was used by survival analysis and related with asymmetrical EP distribution. Many researchers have discussed statistical inference about the parameters in EP distribution using i.i.d random samples. However, sometimes available data might contain only record values, or it is more convenient for researchers to collect record values. We aim to resolve this problem.

We estimated two parameters of the EP distribution by MLE using upper record values. According to simulation study, we used the Bias and MSE of the estimators for studying the efficiency of the proposed estimation method. Then, we discussed the prediction on the next upper record value by known upper record values. The study concluded that MLEs of EP distribution parameters by upper record values has satisfactory performance. Also, prediction of the next upper record value performed well.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. MODEL AND LIKELIHOOD FUNCTION</td>
<td>4</td>
</tr>
<tr>
<td>3. MLE OF SCALE PARAMETER $\lambda$ WITH KNOWN SHAPE PARAMETER</td>
<td>5</td>
</tr>
<tr>
<td>4. MLE OF SHAPE PARAMETER $\theta$ WITH KNOWN SCALE PARAMETER</td>
<td>6</td>
</tr>
<tr>
<td>5. SIMULATIONS FOR MAXIMUM LIKELIHOOD ESTIMATION</td>
<td>11</td>
</tr>
<tr>
<td>6. PREDICTION INTERVAL OF THE NEXT RECORD VALUE</td>
<td>21</td>
</tr>
<tr>
<td>7. CONCLUSION</td>
<td>25</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>26</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bias and MSE from simulations of $\hat{\lambda}$ when $\lambda_0 = 2$ and $\theta_0 = 0.5$</td>
<td>12</td>
</tr>
<tr>
<td>2. Bias and MSE from simulations of $\hat{\theta}$ when $\lambda_0 = 2$ and $\theta_0 = 0.5$</td>
<td>14</td>
</tr>
<tr>
<td>3. Bias and MSE from simulations of $\hat{\lambda}$ and $\hat{\theta}$ for $\lambda_0 = 2$ and $\theta_0 = 0.5$</td>
<td>15</td>
</tr>
<tr>
<td>4. Bias and MSE from simulations of $\hat{\lambda}$ and $\hat{\theta}$ for different $\lambda_0$ and $\theta_0 = 1$</td>
<td>17</td>
</tr>
<tr>
<td>5. Bias and MSE from simulations of $\hat{\lambda}$ and $\hat{\theta}$ for different $\theta_0$ and $\lambda_0 = 1$</td>
<td>18</td>
</tr>
<tr>
<td>6. Average width of prediction interval and coverage probability when $\lambda_0$ is known and $\theta_0$ is unknown</td>
<td>23</td>
</tr>
<tr>
<td>7. Average width of prediction interval and coverage probability when $\lambda_0$ is unknown and $\theta_0$ is known</td>
<td>23</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (a) Bias of $\hat{\lambda}$ (b) MSE of $\hat{\lambda}$ when $\lambda_0 = 2$ and $\theta_0 = 0.5$</td>
<td>13</td>
</tr>
<tr>
<td>2. (a) Bias of $\hat{\theta}$ (b) MSE of $\hat{\theta}$ when $\lambda_0 = 2$ and $\theta_0 = 0.5$</td>
<td>15</td>
</tr>
<tr>
<td>3. (a) Bias of $\hat{\lambda}$ and $\hat{\theta}$ (b) MSE of $\hat{\lambda}$ and $\hat{\theta}$ when $\lambda_0 = 2$ and $\theta_0 = 0.5$</td>
<td>16</td>
</tr>
<tr>
<td>4. (a) Bias of $\hat{\lambda}$ and $\hat{\theta}$ (b) MSE of $\hat{\lambda}$ and $\hat{\theta}$ for different $\lambda_0$ and $\theta_0 = 1$</td>
<td>18</td>
</tr>
<tr>
<td>5. (a) Bias of $\hat{\lambda}$ and $\hat{\theta}$ (b) MSE of $\hat{\lambda}$ and $\hat{\theta}$ for different $\theta_0$ and $\lambda_0 = 1$</td>
<td>19</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

The exponential power (EP) distribution was firstly introduced as a lifetime model by Smith & Bain (1975). The EP distribution has been discussed by many authors, for examples, Leemis (1986), Rajarshi & Rajarshi (1988), Hanagal & Dabade (2015), among others. Moreover, exponential power distribution is not only used by survival analysis but is also related with asymmetrical exponential power distributions in statistics as mentioned in Hazan et al. (2003) and Delicado & Goria (2008).

A random variable is said to have an exponential power distribution with shape parameter $\theta > 0$ and scale parameter $\lambda > 0$ if its probability density function is given by

$$f(t) = \lambda \theta^{\theta - 1} \exp(\lambda t^\theta) \exp(1 - \exp(\lambda t^\theta)), \quad t > 0.$$  

The corresponding survival and hazard rate functions are given by

$$S(t) = P(T > t) = \exp(1 - \exp(\lambda t^\theta)), \quad t > 0$$

and

$$h(t) = \frac{f(x)}{S(x)} = \lambda \theta^{\theta - 1} \exp(\lambda t^\theta), \quad t > 0.$$  

In recent years, many researchers have discussed statistical inference about the parameter in the EP model from different perspectives. These works include Xie et al. (2002), Barriga et al. (2010), Lemonte (2013), etc. All of their estimators of the parameters are estimated by i.i.d. random samples.

However, in many practical situations, either the available data contain only record values, or it is more convenient for researchers to collect record values. We will take care of this concern.
**Definition and real example of record values:**

Let $X_1, X_2, ...$ be an infinite sequence i.i.d. random variables having the same distribution as the population described by random variable $X$. An observation $X_j$ will be called an upper record value (or simply record) if its value exceeds those of all previous observations (Arnolde et al. (1998)). Thus, $X_j$ is a record if $X_j > X_q$, $\forall \ 1 \leq q \leq j-1$. An analogous definition deals with lower record values. Then we assume that $X_j$ is observed at times $j$. The record time sequence $\{T_i, i \geq 0\}$ is defined in the following manner: $T_0 = 1$ with probability 1 and for $i \geq 1$, $T_i = \min \{j : X_j > X_{T_{i-1}}\}$. The record value sequence $\{R_i\}$ is then defined by

$$R_i = X_{T_i}, \quad i = 0, 1, 2, ... .$$

Here $R_0$ is referred to as the reference or the trivial record (Arnolde et al. (1998)).

There are many real examples related with record values. Air quality researchers estimated and obtained confidence interval of parameters of the model for air quality by upper record values (Wu & Tseng (2006), Jafari & Zakerzadeh (2015)). Meanwhile, new joint confidence region for the parameters was obtained using records.

Engineering consideration of the breakdown time of electrical insulating fluid at constant voltage level (Nelson (1982)). Huang & Mi (2015) used records data to compute MLEs of parameters for the model and estimated the prediction interval of the next record value.
The incandescent lamp failure data presented in Davis (1952) was consisted of lifetimes of 417 40-W 110-V incandescent lamps taken from 42 weekly quality control forced-life test samples. Cramer & Naehrig (2012) prompted two versions of generating the record data by the complete sample. They researched MLEs of parameters for the model of incandescent lamp lifetimes using record data.

Therefore, using record values to estimate the parameters of EP distributions will be meaningful and important in those situations. We will investigate the existence and uniqueness of the maximum likelihood estimators of the two parameters $\lambda$ and $\theta$ in the EP distribution using the upper record values. The performance of the MLEs will be explored with simulated data.

The notations, frequently used in the paper are given below for easy reference.

**Notation**

$EPD(\lambda, \theta)$ Exponential power distribution with parameters $\lambda$ and $\theta$.

$\lambda$ Scale parameter in exponential power distribution, $\lambda > 0$

$\theta$ Shape parameter in exponential power distribution, $\theta > 0$

$\{X_i, i \geq 1\}$ i.i.d random variables with $EPD(\lambda, \theta)$

$R_i$ $0 \leq i \leq n$ The first (n+1) upper record values associated with $\{X_i, i \geq 1\}$

$\hat{\lambda}$ Maximum likelihood estimator of $\lambda$

$\hat{\theta}$ Maximum likelihood estimator of $\theta$
2. MODEL AND LIKELIHOOD FUNCTION

Consider exponential power distribution with parameters $\lambda > 0$ and $\theta > 0$. The CDF and pdf are given as

$$F(t) = 1 - \exp(1 - \exp(\lambda t^\theta)), \quad t > 0$$

and

$$f(t) = \lambda \theta t^{\theta-1} \exp(\lambda t^\theta) \exp(1 - \exp(\lambda t^\theta)), \quad t > 0.$$  \hspace{1cm} (1)

The failure rate function of this distribution is

$$h(t) = \frac{f(t)}{S(t)} = \lambda \theta t^{\theta-1} \exp(\lambda t^\theta), \quad t > 0.$$  \hspace{1cm} (2)

Let $R_0, R_1, R_2, \ldots, R_n$ be the first $(n+1)$ upper record values of the exponential power population described by (1)-(3). It is well-known that the joint pdf of $(R_0, R_1, R_2, \ldots, R_n)$ is given by

$$f_{R_0, R_1, \ldots, R_n}(r_0, r_1, \ldots, r_n) = f(r_n) \prod_{i=0}^{n-1} h(r_i), \quad 0 < r_0 < r_1 < \ldots < r_n < \infty.$$ \hspace{1cm} (3)

For details we refer to Arnold (1998) with the help of (4) the likelihood function of parameter $(\lambda, \theta)$ is clearly. Given as

$$L(\theta, \lambda) = (\lambda \theta)^{n+1} \exp(1 - e^{2\pi^2 \theta}) \prod_{i=0}^{n-1} r_i^{\theta-1} e^{2\pi^2 \theta}.$$ \hspace{1cm} (4)

Hence the log-likelihood function of $(\lambda, \theta)$ is

$$l(\theta, \lambda) = \ln L(\theta, \lambda) = (n+1) \ln(\lambda) + (n+1) \ln(\theta) + (1 - e^{2\pi^2 \theta}) + (\theta - 1) \sum_{i=0}^{n} \ln(r_i) + \lambda \sum_{i=0}^{n} r_i^{\theta}.$$ \hspace{1cm} (5)
3. MLE OF SCALE PARAMETER $\lambda$ WITH KNOWN SHAPE PARAMETER

In this section we study the MLE of parameter $\lambda$. Assume $\theta = \theta_0$ is known.

From (6) we donate

$$l_i(\lambda) = (n + 1) \ln(\lambda) + (n + 1) \ln(\theta_0) + (1 - e^{\lambda n \theta_0}) + (\theta_0 - 1) \sum_{i=0}^{n} \ln(r_i) + \lambda \sum_{i=0}^{n} r_i^{\theta_0} \quad (7)$$

We have

$$l'_i(\lambda) = \frac{dl_i(\lambda)}{d\lambda} = \frac{n + 1}{\lambda} - e^{\lambda n \theta_0} r_n^{\theta_0} + \sum_{i=0}^{n} r_i^{\theta_0} \quad (8)$$

It is easy to see that

$$\lim_{\lambda \to 0^+} l'_i(\lambda) = +\infty \quad (9)$$

and

$$\lim_{\lambda \to +\infty} l'_i(\lambda) = -\infty \quad (10)$$

That imply that the likelihood equation $l'_i(\lambda) = 0$ has at least one solution in $(0, \infty)$.

We further have

$$l''_i(\lambda) = \frac{d^2l_i(\lambda)}{d\lambda^2} = -\frac{n + 1}{\lambda^2} - e^{\lambda n \theta_0} r_n^{2\theta_0} < 0$$

Therefore, $l'_i(\lambda)$ is a concave function of $\lambda \in (0, \infty)$. The above results show that the log-likelihood function $l_i(\lambda)$ attain its maximum over $(0, \infty)$ at a unique point $\hat{\lambda}$.

Therefore, the MLE of $\lambda$ uniquely exacts and is the unique solution of the equation:

$$\frac{n + 1}{\lambda} - e^{\lambda n \theta_0} r_n^{\theta_0} + \sum_{i=0}^{n} r_i^{\theta_0} = 0 \quad (11)$$
4. MLE OF SHAPE PARAMETER $\Theta$ WITH KNOWN SCALE PARAMETER

Let the scale parameter $\lambda = \lambda_0$ is known. In this case, we express the log-likelihood function of $\theta$ as:

$$l_z(\theta) = (n+1) \ln(\lambda_0) + (n+1) \ln(\theta) + (1 - e^{\lambda_0 \theta}) + (\theta - 1) \sum_{i=0}^{n} \ln(r_i) + \lambda_0 \sum_{i=0}^{n} r_i^\theta.$$  \hspace{1cm}(12)

We have

$$l_z'(\theta) = \frac{d l_z(\theta)}{d \theta} = \frac{n+1}{\theta} - \lambda_0 e^{\lambda_0 \theta} r_n^\theta \ln(r_n) + \sum_{i=0}^{n} \ln(r_i) + \lambda_0 \sum_{i=0}^{n} r_i^\theta \ln(r_i).$$  \hspace{1cm}(13)

Below we explore the limiting behavior of $l_z'(\theta)$ as $\theta \to 0^+$ and $\theta \to +\infty$.

Claim 1.

$$\lim_{\theta \to 0^+} l_z'(\theta) = +\infty.$$  \hspace{1cm}(14)

Since when $\theta \to 0^+$,

$$\frac{n+1}{\theta} \to +\infty$$

and

$$\lambda_0 e^{\lambda_0 \theta} r_n^\theta \ln(r_n) + \sum_{i=0}^{n} \ln(r_i) + \lambda_0 \sum_{i=0}^{n} r_i^\theta \ln(r_i) \to \lambda_0 e^{\lambda_0 \theta} \ln(r_n) + \sum_{i=0}^{n} \ln(r_i) + \lambda_0 \sum_{i=0}^{n} \ln(r_i) > -\infty$$

Claim 2.

$$\lim_{\theta \to +\infty} l_z'(\theta) < 0.$$  \hspace{1cm}(15)

As $\theta \to +\infty$ it holds that

$$\lim_{\theta \to +\infty} l_z'(\theta) = \sum_{i=0}^{n} \ln(r_i) - \lambda_0 \lim_{\theta \to +\infty}[r_n^\theta \ln(r_n)e^{\lambda_0 \theta} - \sum_{i=0}^{n} r_i^\theta \ln(r_i)].$$  \hspace{1cm}(16)
Since \( \frac{n+1}{\theta} \to 0 \), and
\[
\sum_{i=0}^{n} \ln(r_i) - \lambda_n r_n^\theta \ln(r_n)[e^{\lambda_n r_n^\theta} - 1]
\]
when \( \theta \to +\infty \),
\[
\frac{r_i^\theta \ln(r_i)}{r_n^\theta \ln(r_n)} = \left( \frac{r_i}{r_n} \right)^\theta \left( \frac{\ln(r_i)}{\ln(r_n)} \right) \to 0,
\]
and here
\[
\lim \sum_{i=0}^{n} \frac{r_i^\theta \ln(r_i)}{r_n^\theta \ln(r_n)} = 0 + 0 + \cdots + \left( \frac{r_n}{r_n} \right)^\theta \left( \frac{\ln(r_n)}{\ln(r_n)} \right) = 1.
\]
Thus,
\[
\lim l_i'(\theta) = \sum_{i=0}^{n} \ln(r_i) - \lambda_n \lim r_n^\theta \ln(r_n)[e^{\lambda_n r_n^\theta} - 1].
\]

Now we study 3 case separately.

Case 1.

\( r_n > 1 \).

In this case, clearly
\[
\lim l_i'(\theta) = -\infty.
\]

Case 2.

\( r_n < 1 \).

We have \( r_n^\theta \to 0 \) and hence \( r_n^\theta \ln(r_n)[e^{\lambda_n r_n^\theta} - 1] \to 0 \).

There functions imply
\[
\lim l_i'(\theta) = \sum_{i=0}^{n} \ln(r_i) < 0 \text{ since } 0 < r_0 < r_1 < \cdots < r_n < 1.
\]
Case 3.

\[ r_n = 1. \]

We have \( r_n^0 = 1. \)

\[ \lim_{\theta \to +\infty} l'_2(\theta) = \sum_{i=0}^{n} \ln(r_i) - \lambda_0 \ln(1)(e^{\lambda_0} - 1) < 0 \]

since \( 0 < r_0 < r_1 < \ldots < r_{n-1} < 1 \) and \( \lambda_0 > 0. \)

The above argument show that at any case

\[ \lim_{\theta \to +\infty} l'_2(\theta) < 0. \]

We have seem that \( \lim_{\theta \to +\infty} l'_2(\theta) = +\infty \) and \( \lim_{\theta \to +\infty} l'_2(\theta) < 0 \) so that equation \( l'_2(\theta) = 0 \)

has at least one solution in the interval \((0, \infty)\).

Recall that \( l_2(\theta) \) is given in (12) we have

\[ \lim_{\theta \to 0^+} l_2(\theta) = -\infty. \]

In order to investigate the sign of \( \lim_{\theta \to +\infty} l_2(\theta) \), we reconsider the above three case.

Case 1.

\[ r_n > 1. \]

In this case, we have three facts:

Fact 1: \( \lim_{\theta \to +\infty} \frac{r_i^\theta}{e^{\lambda_0 r_i^\theta}} = 0. \)

Fact 2: \( \lim_{\theta \to +\infty} \frac{\theta - 1}{e^{\lambda_0 \theta}} = 0. \)

Fact 3: \( \lim_{\theta \to +\infty} \frac{\ln(\theta)}{e^{\lambda_0 \theta}} = 0. \)
Then

\[
\lim_{\theta \to +\infty} l_2(\theta) = 1 + (n + 1) \ln(\lambda_0) + \lim_{\theta \to +\infty} e^{\lambda_0 r_\theta} \left[ \frac{(n + 1) \ln(\theta)}{e^{\lambda_0 r_\theta}} + \frac{\theta - 1}{\theta} \sum_{i=0}^{\infty} \ln(r_i) + \sum_{i=0}^{\infty} \frac{r_i^\theta}{\theta} - 1 \right]
\]

According above three facts, we have

\[
\lim_{\theta \to +\infty} l_2(\theta) = 1 + (n + 1) \ln(\lambda_0) + \lim_{\theta \to +\infty} e^{\lambda_0 r_\theta} (-1) = -\infty .
\]

Case 2.

\[
0 < r_n < 1 .
\]

In this case,

\[
\lim_{\theta \to +\infty} l_2(\theta) = (n + 1) \ln(\lambda_0) + (1 - e^{\lambda_0 r_\theta}) + \lim_{\theta \to +\infty} (n + 1) \ln(\theta) \left[ 1 + \frac{\theta - 1}{(n + 1) \ln(\theta)} \sum_{i=0}^{\infty} \ln(r_i) \right] - (\theta - 1) \sum_{i=0}^{\infty} \ln(r_i)
\]

\[
= -\infty
\]

Case 3.

\[
r_n = 1 .
\]

\[
\lim_{\theta \to +\infty} l_2(\theta) = (n + 1) \ln(\lambda_0) + (1 - e^{\lambda_0}) + \lim_{\theta \to +\infty} (n + 1) \ln(\theta) \left[ 1 + \frac{\theta - 1}{(n + 1) \ln(\theta)} \sum_{i=0}^{\infty} \ln(r_i) \right] + \lambda_0 \lim_{\theta \to +\infty} \sum_{i=0}^{\infty} r_i^\theta
\]

\[
= -\infty .
\]

Summarizing the above, we have shown that 

\[
\lim_{\theta \to 0^+} l_2(\theta) = \lim_{\theta \to +\infty} l_2(\theta) = -\infty
\]

and the equation:

\[
l_2'(\theta) = 0
\]

(17)

has at least one solution in \((0, \infty)\). Therefore, \(l_2(\theta)\) a continuous function of \(\theta\) on \((0, \infty)\) must attain it maximum at some interior point of \((0, \infty)\).
Unfortunately the uniqueness of the MLE of $\theta$ is still an open question. However, we can compute the MLE $\hat{\theta}$ of $\theta$.

All in all, in order to estimate $\theta$ and $\lambda$, we need to solve the system of equations

\[ \frac{n+1}{\lambda} - e^{\frac{1}{\lambda}} r_n^\theta + \sum_{i=0}^{n} r_i^\theta = 0 \]

\[ \frac{n+1}{\theta} - \lambda e^{\frac{1}{\theta}} r_n^\theta \ln(r_n) + \sum_{i=0}^{n} \ln(r_i) + \lambda \sum_{i=0}^{n} r_i^\theta \ln(r_i) = 0. \quad (18) \]

Since this non-linear equation set cannot be solved directly, a numerical root finding technique must be used. There are a lot of ways can be used to find the roots. We will use the well-known Newton-Raphson method.
5. SIMULATIONS FOR MAXIMUM LIKELIHOOD ESTIMATION

We describe the computer simulations and discuss the behavior of the maximum likelihood estimators of the exponential power distribution parameters from upper record data. The analytical work has been done by using R.

**Step 1: Generated** $x_0, x_1, x_2, \ldots$ **from exponential power distribution with parameters** $\lambda$ **and** $\theta$.

To this purpose, the probability integrate transform is employed. Thus, we first generate $u_0, u_1, u_2, \ldots$ from Uniform distribution $(0,1)$. Then, solve the following equation for $x$.

$$u = F(x) = 1 - \exp(1 - e^{\lambda x^\theta}) .$$

We can get

$$x = F^{-1}(u) = \left(\frac{\ln(1 - \ln(1 - u))}{\lambda}\right)^{1/\theta} .$$

By this way, we can produce random observation $x_0, x_1, x_2, \ldots$ from R project easily.

**Step 2: Obtain a sample of upper record value.**

A record sample $r_0, r_1, r_2, \ldots$ was produced form sequence $x_0, x_1, x_2, \ldots$ obtained in step 1. R code upper.record.value can help us to do it.

**Step 3: Calculate the MLEs** $\hat{\lambda}$ **and** $\hat{\theta}$.

The MLEs of $\hat{\lambda}$ and $\hat{\theta}$ for the record sample are found by Newton-Raphson method using equation (18). Name the resulting estimates as $\hat{\lambda}_i$ and $\hat{\theta}_i$. 
Step 4: Repeat step 1-3 1000 times

The above process repeated 1000 times. Consequently, we have a set of 1000 parameter estimates using method.

Step 5: Evaluate the performance of $\hat{\lambda}$ and $\hat{\theta}$.

For each pair of given $(\lambda, \theta)$. Compute the mean, bias and mean squared error (MSE) for each of estimates of the parameters using the 1000 values of $\hat{\lambda}$ and $\hat{\theta}$.

The procedure described above was repeated for different complete sample with sizes $m=50, 100, 200, 500, 1000, 2000, 5000, 10000$. Then the mean values $\bar{\lambda}$ and $\bar{\theta}$ and mean squared errors values $MSE_\lambda$ and $MSE_\theta$ were computed using:

\[
\bar{\lambda} = \frac{1}{1000} \sum_{i=1}^{1000} \hat{\lambda}_i , \quad \bar{\theta} = \frac{1}{1000} \sum_{i=1}^{1000} \hat{\theta}_i ,
\]

\[
MSE_\lambda = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{\lambda}_i - \lambda_0)^2 , \quad MSE_\theta = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{\theta}_i - \theta_0)^2.
\]

Firstly we study the behavior of $\hat{\lambda}$. We assume $\lambda_0 = 2$ and $\theta_0 = 0.5$. Table 1 and Figure 1 show the simulation results of $\hat{\lambda}$.

<table>
<thead>
<tr>
<th>m</th>
<th>Bias of $\hat{\lambda}$</th>
<th>MSE of $\hat{\lambda}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>59.0072</td>
<td>108027.6159</td>
</tr>
<tr>
<td>100</td>
<td>31.1752</td>
<td>67270.6156</td>
</tr>
<tr>
<td>200</td>
<td>6.9452</td>
<td>10458.8708</td>
</tr>
<tr>
<td>500</td>
<td>0.0313</td>
<td>0.2179</td>
</tr>
<tr>
<td>1000</td>
<td>-0.0458</td>
<td>0.1553</td>
</tr>
<tr>
<td>2000</td>
<td>-0.0679</td>
<td>0.1272</td>
</tr>
<tr>
<td>5000</td>
<td>-0.0968</td>
<td>0.1351</td>
</tr>
<tr>
<td>10000</td>
<td>-0.0845</td>
<td>0.1137</td>
</tr>
</tbody>
</table>
According to Table 1 and Figure 1, we could see that the absolute value of the Bias and MSE of $\hat{\lambda}$ decreases from $m=50$ to $m=10000$ clearly. When $m$ is larger than 500, both bias and MSE decrease quickly. Since we could get enough number of upper record values from random sample when complete sample sizes are greater than 500. In this
situation, estimating \( \lambda \) by record values is reasonable. Secondly, we study the behavior of \( \hat{\theta} \). We assume \( \lambda_0 = 2 \) and \( \theta_0 = 0.5 \). Table 2 and Figure 2 show the simulation results of \( \hat{\theta} \).

<table>
<thead>
<tr>
<th>m</th>
<th>Bias of ( \hat{\theta} )</th>
<th>MSE of ( \hat{\theta} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>1.892</td>
<td>24.6591</td>
</tr>
<tr>
<td>100</td>
<td>1.5346</td>
<td>24</td>
</tr>
<tr>
<td>200</td>
<td>0.9354</td>
<td>13.3547</td>
</tr>
<tr>
<td>500</td>
<td>0.5441</td>
<td>3.1677</td>
</tr>
<tr>
<td>1000</td>
<td>0.4482</td>
<td>2.7563</td>
</tr>
<tr>
<td>2000</td>
<td>0.3425</td>
<td>1.4081</td>
</tr>
<tr>
<td>5000</td>
<td>0.2332</td>
<td>0.3074</td>
</tr>
<tr>
<td>10000</td>
<td>0.1963</td>
<td>0.2692</td>
</tr>
</tbody>
</table>

![Bias and MSE from simulations of \( \hat{\theta} \) when \( \lambda_0 = 2 \) and \( \theta_0 = 0.5 \).](image-url)
We observe that bias and MSE of \( \hat{\theta} \) decreases from \( m=50 \) to \( m=10000 \) by Table 2 and Figure 2 explicitly. When \( m \) is larger than 500, both bias and MSE decrease quickly. The reason is as similar as the case of \( \hat{\lambda} \).

Next the procedure described above was repeated for different sample sizes of upper record values \( k=5, 6, 7, 8, 9, 10 \). We assume \( \lambda_0 = 2 \) and \( \theta_0 = 0.5 \). We have got Table 3 and Figure 3.

<table>
<thead>
<tr>
<th>k</th>
<th>Bias of ( \hat{\lambda} )</th>
<th>MSE of ( \hat{\lambda} )</th>
<th>Bias of ( \hat{\theta} )</th>
<th>MSE of ( \hat{\theta} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.4664</td>
<td>91.3631</td>
<td>1.0556</td>
<td>2.5865</td>
</tr>
<tr>
<td>5</td>
<td>-0.6896</td>
<td>0.5673</td>
<td>0.7084</td>
<td>1.1334</td>
</tr>
<tr>
<td>6</td>
<td>-0.5081</td>
<td>0.3178</td>
<td>0.4855</td>
<td>0.5172</td>
</tr>
<tr>
<td>7</td>
<td>-0.3857</td>
<td>0.185</td>
<td>0.339</td>
<td>0.2836</td>
</tr>
<tr>
<td>8</td>
<td>-0.2545</td>
<td>0.0947</td>
<td>0.247</td>
<td>0.1864</td>
</tr>
<tr>
<td>9</td>
<td>-0.1475</td>
<td>0.0575</td>
<td>0.1803</td>
<td>0.0989</td>
</tr>
<tr>
<td>10</td>
<td>-0.0399</td>
<td>0.0244</td>
<td>0.1457</td>
<td>0.0822</td>
</tr>
</tbody>
</table>
Figure 3 (a) Bias of $\hat{\lambda}$ and $\hat{\theta}$ (b) MSE of $\hat{\lambda}$ and $\hat{\theta}$ when $\lambda_0 = 2$ and $\theta_0 = 0.5$.

According to Table 3 and Figure 3, We can see that there are good behavior of $\hat{\lambda}$ and $\hat{\theta}$ when $k \geq 5$.

Then we have researched behavior of $\hat{\lambda}$ when complete sample size $m=10000$ and $\lambda_0 = 0.25, 0.5, 1, 2, 4, 8, 16, 32$ and $\theta_0 = 1$. We have get Table 4 and Figure 4.
Table 4. Bias and MSE from simulations of $\hat{\lambda}$ and $\hat{\theta}$ for different $\lambda_0$ and $\theta_0 = 1$.

<table>
<thead>
<tr>
<th>$\lambda_0$</th>
<th>Bias of $\hat{\lambda}$</th>
<th>MSE of $\hat{\lambda}$</th>
<th>Bias of $\hat{\theta}$</th>
<th>MSE of $\hat{\theta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.1008</td>
<td>0.0438</td>
<td>-0.0732</td>
<td>0.9378</td>
</tr>
<tr>
<td>0.5</td>
<td>0.07</td>
<td>0.4239</td>
<td>0.0282</td>
<td>0.8841</td>
</tr>
<tr>
<td>1</td>
<td>-0.0705</td>
<td>0.1155</td>
<td>0.1871</td>
<td>0.2545</td>
</tr>
<tr>
<td>2</td>
<td>-0.1083</td>
<td>0.1425</td>
<td>0.4101</td>
<td>0.8475</td>
</tr>
<tr>
<td>4</td>
<td>6.7628</td>
<td>668.7936</td>
<td>0.835</td>
<td>3.8634</td>
</tr>
<tr>
<td>8</td>
<td>21.0445</td>
<td>1231.2698</td>
<td>0.7886</td>
<td>0.7549</td>
</tr>
<tr>
<td>16</td>
<td>30.0993</td>
<td>1455.9522</td>
<td>0.502</td>
<td>0.9023</td>
</tr>
<tr>
<td>32</td>
<td>29.9711</td>
<td>1342.8429</td>
<td>0.2458</td>
<td>0.9763</td>
</tr>
</tbody>
</table>
Figure 4 (a) Bias of $\hat{\lambda}$ and $\hat{\Theta}$ (b) MSE of $\hat{\lambda}$ and $\hat{\Theta}$ for different $\lambda_0$ and $\theta_0 = 1$.

According to Table 4 and Figure 4. We can conclude that with the increase of $\lambda_0$ bias and MSE of $\hat{\lambda}$ increase. Bias and MSE of $\hat{\Theta}$ do not have great change. The unexpected tendency of $\hat{\lambda}$ suggests that maximum likelihood estimation of parameter $\lambda$ of the exponential power distribution from record samples is not effective when $\lambda$ is too large. However, as $\lambda_0$ decreases, the MSE of $\hat{\lambda}$ decreases.

Finally, we study behavior of $\hat{\Theta}$ when complete sample size $m=10000$ and $\lambda_0 = 1$ and $\theta_0 = 0.25, 0.5, 1, 2, 4, 8, 16, 32$. We have get Table 5 and Figure 5.

<table>
<thead>
<tr>
<th>$\theta_0$</th>
<th>Bias of $\hat{\lambda}$</th>
<th>MSE of $\hat{\lambda}$</th>
<th>Bias of $\hat{\Theta}$</th>
<th>MSE of $\hat{\Theta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>-0.4367</td>
<td>0.3757</td>
<td>0.3006</td>
<td>0.2122</td>
</tr>
<tr>
<td>0.5</td>
<td>-0.0621</td>
<td>0.2021</td>
<td>0.0856</td>
<td>0.1054</td>
</tr>
<tr>
<td>1</td>
<td>-0.0705</td>
<td>0.1155</td>
<td>0.1871</td>
<td>0.2545</td>
</tr>
<tr>
<td>2</td>
<td>-0.4405</td>
<td>0.2528</td>
<td>0.5508</td>
<td>0.433</td>
</tr>
<tr>
<td>4</td>
<td>-0.0637</td>
<td>0.1223</td>
<td>0.7096</td>
<td>3.9918</td>
</tr>
<tr>
<td>8</td>
<td>-0.0569</td>
<td>0.1168</td>
<td>0.6428</td>
<td>3.7206</td>
</tr>
</tbody>
</table>
Figure 5 (a) Bias of $\hat{\lambda}$ and $\hat{\theta}$ (b) MSE of $\hat{\lambda}$ and $\hat{\theta}$ for different $\theta_0$ and $\lambda_0 = 1$. 

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>-0.0532</td>
<td>0.1272</td>
<td>1.3168</td>
<td>15.492</td>
</tr>
<tr>
<td>32</td>
<td>-0.0365</td>
<td>0.1085</td>
<td>4.4169</td>
<td>186.62</td>
</tr>
</tbody>
</table>
For samples of records, the MSEs of $\hat{\lambda}$ do not have great change with increase of $\theta$. In the general, MSEs of $\hat{\theta}$ increase slightly.

All in all, when $\lambda_0$ and $\theta_0$ are not every large, the estimates of $\lambda$ and $\theta$ have a good performance. Meanwhile, with the increase of the complete sample size $m$ or the record sample size $k$, bias and MSE of $\hat{\lambda}$ and $\hat{\theta}$ decrease quickly.
6. PREDICTION INTERVAL OF THE NEXT RECORD VALUE

Assume that we have observed the first (n+1) upper record values \( r_0, r_1, r_2, \ldots, r_n \) from the exponential power distribution. Then we can make certain prediction on the next record value \( R_{n+1} \).

In order to be able to make prediction on \( R_{n+1} \) using \( \mathbf{r} \equiv (r_0, r_1, r_2, \ldots, r_n) \) we need the conditional density of \( R_{n+1} \) given \( \mathbf{r} \). The conditional distribution of \( R_{n+1} \) given \( \mathbf{r} \) is obtained as

\[
 f_{R_{n+1} | \mathbf{r}}(r_{n+1} | \mathbf{r}) = \frac{f(r_0, \ldots, r_n, r_{n+1})}{f(r_0, \ldots, r_n)} = \frac{f(r_{n+1}) \prod_{i=0}^{n} h(r_i)}{f(r_n) \prod_{i=0}^{n-1} h(r_i)}
\]

\[
 = \frac{f(r_{n+1}) h(r_n)}{f(r_n)} = \frac{f(r_{n+1}) h(r_n)}{F(r_n) h(r_n)}
\]

\[
 = \frac{f(r_{n+1})}{F(r_n)} \quad r_{n+1} > r_n.
\] (21)

Let the density function of random variable \( W \) be given by (21). Obviously, \( W \) has survival function

\[
 \bar{G}(w) = P(W > w) = P(R_{n+1} > w | \mathbf{r}) = \int_{w}^{\infty} f_{R_{n+1} | \mathbf{r}}(r_{n+1} | \mathbf{r}) dr_{n+1}
\]

\[
 = \int_{w}^{\infty} \frac{f(r_{n+1})}{F(r_n)} dr_{n+1} = \frac{\bar{F}(w)}{F(r_n)}
\]

\[
 = \frac{\exp(1 - \exp(\lambda w^\theta))}{\exp(1 - \exp(\lambda r_n^\theta))}
\]

\[
 = \frac{\exp(-\exp(\lambda w^\theta))}{\exp(-\exp(\lambda r_n^\theta))}.
\] (22)
For any given \( \delta \in (0,1) \) let \( c_\delta \) be the \( \delta \)th quantile of \( G(w) \). That imply \( \overline{G}(c_\delta) = \delta \).

The value \( c_\delta \) can be obtained by solving the equation
\[
\frac{\exp(-\exp(\lambda c_\delta)})}{\exp(-\exp(\lambda r_\alpha^\theta))} = \delta.
\]

We obtained
\[
c_\delta = \left( \frac{1}{\hat{\lambda}} \ln[\exp(\lambda r_\alpha^\theta) - \ln(\delta)] \right)^{1/\theta}.
\]

Assuming we have \( \theta_0 \). Substituting \( \hat{\lambda} \) and \( \theta_0 \) for \( \lambda \) and \( \theta \) in (23) respectively, we can have the approximate value of \( c_\delta \).
\[
c_\delta \approx \hat{c}_\delta = \left( \frac{1}{\hat{\lambda}} \ln[\exp(\hat{\lambda} r_\alpha^\theta) - \ln(\delta)] \right)^{1/\theta_0}.
\]

Similarly, we assume we have \( \lambda_0 \). Then substituting \( \lambda_0 \) and \( \hat{\theta} \) for \( \lambda \) and \( \theta \) in (23) respectively, we can have the approximate value of \( c_\delta \).
\[
c_\delta \approx \hat{c}_\delta = \left( \frac{1}{\lambda_0} \ln[\exp(\lambda_0 r_\alpha^\theta) - \ln(\delta)] \right)^{1/\hat{\theta}}.
\]

Now, for any given \( \alpha \in (0,1) \), a \((1-\alpha)\) prediction interval of \( R_{n+1} \) given \( r \) can expressed as
\[
(\hat{c}_{1-\alpha/2}, \hat{c}_{\alpha/2}).
\]

**Simulation Study**

We first generated \( x_0, x_1, x_2, \ldots \) from exponential power distribution with parameters \( \lambda_0 \) and \( \theta_0 \). We took 10 upper record values \( r = (r_0, r_1, \ldots, r_9) \) and \( r_\alpha \) from the sequence...
Then we computed $90\%$ prediction interval of $R_g$ by $r$ from two situations. With $\lambda_0$ known and $\theta_0$ unknown, we calculated $\hat{c}_g$ from (24). With $\lambda_0$ unknown and $\theta_0$ known, we calculated $\hat{c}_g$ from (25). Repeating the above steps 1000 times, we got the average width of the prediction interval and the coverage probability $P(r_0 \in (\hat{c}_{1-\alpha/2}, \hat{c}_{\alpha/2}))$.

The results were shown as Table 6 and Table 7.

<table>
<thead>
<tr>
<th>$\lambda_0$</th>
<th>$\theta_0$</th>
<th>Average Width</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>4.3346</td>
<td>0.747</td>
</tr>
<tr>
<td>0.5</td>
<td>0.75</td>
<td>1.0385</td>
<td>0.766</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>0.4636</td>
<td>0.745</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>1.0915</td>
<td>0.739</td>
</tr>
<tr>
<td>1</td>
<td>0.75</td>
<td>0.4116</td>
<td>0.735</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.2332</td>
<td>0.762</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>0.2718</td>
<td>0.741</td>
</tr>
<tr>
<td>2</td>
<td>0.75</td>
<td>0.1633</td>
<td>0.763</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.1158</td>
<td>0.722</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\lambda_0$</th>
<th>$\theta_0$</th>
<th>Average Width</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>4.3343</td>
<td>0.734</td>
</tr>
<tr>
<td>0.5</td>
<td>0.75</td>
<td>1.0293</td>
<td>0.736</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>0.4633</td>
<td>0.725</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>1.0493</td>
<td>0.72</td>
</tr>
<tr>
<td>1</td>
<td>0.75</td>
<td>0.3987</td>
<td>0.733</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.2253</td>
<td>0.706</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>0.2554</td>
<td>0.683</td>
</tr>
<tr>
<td>2</td>
<td>0.75</td>
<td>0.1489</td>
<td>0.713</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.1059</td>
<td>0.705</td>
</tr>
</tbody>
</table>

According to Table 6 and Table 7, the performance of the prediction intervals $R_g$ by $r$ is effective. Thus, it is reasonable for us to apply the derived prediction interval for $x_0, x_1, x_2, \ldots$. Then we computed $90\%$ prediction interval of $R_g$ by $r$ from two situations. With $\lambda_0$ known and $\theta_0$ unknown, we calculated $\hat{c}_g$ from (24). With $\lambda_0$ unknown and $\theta_0$ known, we calculated $\hat{c}_g$ from (25). Repeating the above steps 1000 times, we got the average width of the prediction interval and the coverage probability $P(r_0 \in (\hat{c}_{1-\alpha/2}, \hat{c}_{\alpha/2}))$.

The results were shown as Table 6 and Table 7.

<table>
<thead>
<tr>
<th>$\lambda_0$</th>
<th>$\theta_0$</th>
<th>Average Width</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>4.3346</td>
<td>0.747</td>
</tr>
<tr>
<td>0.5</td>
<td>0.75</td>
<td>1.0385</td>
<td>0.766</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>0.4636</td>
<td>0.745</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>1.0915</td>
<td>0.739</td>
</tr>
<tr>
<td>1</td>
<td>0.75</td>
<td>0.4116</td>
<td>0.735</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.2332</td>
<td>0.762</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>0.2718</td>
<td>0.741</td>
</tr>
<tr>
<td>2</td>
<td>0.75</td>
<td>0.1633</td>
<td>0.763</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.1158</td>
<td>0.722</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\lambda_0$</th>
<th>$\theta_0$</th>
<th>Average Width</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>4.3343</td>
<td>0.734</td>
</tr>
<tr>
<td>0.5</td>
<td>0.75</td>
<td>1.0293</td>
<td>0.736</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>0.4633</td>
<td>0.725</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>1.0493</td>
<td>0.72</td>
</tr>
<tr>
<td>1</td>
<td>0.75</td>
<td>0.3987</td>
<td>0.733</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.2253</td>
<td>0.706</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>0.2554</td>
<td>0.683</td>
</tr>
<tr>
<td>2</td>
<td>0.75</td>
<td>0.1489</td>
<td>0.713</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.1059</td>
<td>0.705</td>
</tr>
</tbody>
</table>
predicting the value of the \((n+1)^{th}\) upper record by \(\mathbf{r} = (r_0, r_1, \ldots, r_n)\) from (24), (25) and (26).
7. CONCLUSION

In the study presented above, we researched MLEs of exponential power distribution parameters by upper record values and discussed the uniqueness of MLEs. We then used simulation study to research the performance of $\hat{\lambda}$ and $\hat{\theta}$ and concluded that they have good performance in most situations. Finally, we studied the prediction of $R_{n+1}$ from $r$. According to the simulation study, we concluded that using $r$ to estimate $R_{n+1}$ is reasonable.
REFERENCE


