Research Article

Exact Interval Inference for the Two-Parameter Rayleigh Distribution Based on the Upper Record Values

Jung-In Seo,¹ Jae-Woo Jeon,² and Suk-Bok Kang²

¹Department of Statistics, Daejeon University, No. 62, Daehak-ro, Dong-gu, Republic of Korea
²Department of Statistics, Yeungnam University, No. 280, Daehak-ro, Gyeongsan, Republic of Korea

Correspondence should be addressed to Suk-Bok Kang; sbkang@yu.ac.kr

Received 22 July 2016; Revised 20 October 2016; Accepted 8 November 2016

Academic Editor: Shesh N. Rai

Copyright © 2016 Jung-In Seo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The maximum likelihood method is the most widely used estimation method. On the other hand, it can produce substantial bias, and an approximate confidence interval based on the maximum likelihood estimator cannot be valid when the sample size is small. Because the sizes of the record values are considerably smaller than the original sequence observed in the majority of cases, a method appropriate for this situation is required for precise inference. This paper provides the exact confidence intervals for unknown parameters and exact predictive intervals for the future upper record values by providing some pivotal quantities in the two-parameter Rayleigh distribution based on the upper record values. Finally, the validity of the proposed inference methods was examined from Monte Carlo simulations and real data.

1. Introduction

The cumulative distribution function (cdf) and probability density function (pdf) of the random variable (RV), \( X \), with the Rayleigh distribution are given, respectively, by

\[
F(x) = 1 - \exp \left[ -\frac{(x - \mu)^2}{2\sigma^2} \right],
\]

\[
f(x) = \frac{x - \mu}{\sigma^2} \exp \left[ -\frac{(x - \mu)^2}{2\sigma^2} \right], \quad x > \mu, \quad \sigma > 0,
\]

where \( \mu \) is the location parameter and \( \sigma \) is the scale parameter. The Rayleigh distribution is used because the life of the model theory reliability plays an important role in modeling the life of the random phenomenon. Moreover, it is used in many applications, such as reliability, life tests, and survival analysis because its failure rate is a linear function of time. Therefore, this distribution has been studied by many authors in the case, where samples are censored due to a range of reasons. Dyer and Whisendan [1] examined the properties of the \( k \)-optimum best linear unbiased estimators (BLUEs) of the scale parameter in the Rayleigh distribution and provided an approximate \( k \)-optimum BLUE based on \( k \) order statistics. Sinha and Howlader [2] derived the highest posterior density (HPD) credible interval for the scale parameter and the reliability function in a Rayleigh distribution. Ali Mousa and Al-Sagheer [3] obtained the maximum likelihood estimators (MLEs) and Bayes estimators for \( \mu, \sigma \), and the reliability function of the Rayleigh distribution obtained based on the progressively Type-II censored data. Raqab and Madi [4] discussed the Bayesian predictive methods for the total time on test using doubly censored data with a Rayleigh distribution and the scale parameter and applied the methods to a real data set that represented the deep-groove ball bearing failure times. For the same real data, Kim and Han [5] applied a Bayesian inference method based on the conjugate prior of the scale parameter of the Rayleigh distribution under general progressive censoring and S. Dey and T. Dey [6] applied this by providing point and interval estimation methods for the scale parameter of the Rayleigh distribution under progressive Type-II censoring with binomial removal. This paper considered a two-parameter Rayleigh distribution based on the upper record values that are used extensively to build statistical modeling arising in many real-life situations involving weather, sports, economics, and life tests. The record values are described as follows.
Let \( \{X_i, i = 1, \ldots, n\} \) be a sequence of independent and identically distributed (i.i.d) RVs from a continuous probability distribution. If \( X_j > X_i \) for all \( i < j \), then \( X_j \) is an upper record value. The indices at which the upper record values occur are given by the record times \( \{U(k), k \geq 1\} \), where \( U(k) = \min\{j \mid j > U(k-1), X_j > X_{U(k-1)}\} \), \( k \geq 1 \) with \( U(1) = 1 \). Chandler [7] first studied the record values and their basic properties. Ahsanullah [8] provided detailed descriptions of the general theory and applications for the well-known probability distributions based on the records. Seo and Kim [9] provided inference methods to estimate unknown parameters and predicted future upper record values from the extreme value distribution using both frequentist and Bayesian approaches. Note that the sizes of the record values are actually considerably smaller than the observed original sequence in the majority of case; a method appropriate for this situation is required for precise inference. The maximum likelihood method is the most extensively used estimation method. On the other hand, the approximate confidence intervals (CIs) based on the asymptotic normality of the MLE can yield inappropriate results because \( \mu \) and \( \sigma \) are supported by \((-\infty, x)\) and \((0, \infty)\), respectively. Moreover, the asymptotic normality of the MLE requires the suitable regularity conditions and it is difficult to prove that the regularity conditions are satisfied when the record values are observed from the two-parameter Rayleigh distribution. This paper constructs exact CIs for unknown parameters \((\mu, \sigma)\) of the Rayleigh distribution based on the upper record values by providing some pivotal quantities, which are much more efficient than the maximum likelihood method in terms of computation cost. Another aim of this paper is to construct exact predictive intervals (PIs) for the future upper record values based on the past upper record values from the Rayleigh distribution because it is very important to correctly predict in many fields such as earthquakes, flood, and rainfall.

The remainder of the paper is structured as follows. Section 2 provides some pivotal quantities and derives the exact CIs for unknown parameters and PIs for the future upper record values in the Rayleigh distribution based on the upper record values. Section 3 assesses the validity of the proposed method through Monte Carlo simulations and real data. Section 4 concludes the paper.

### 2. Inference Based on Pivotal Quantity

The likelihood function for \( \theta \) is given by Arnold et al., [10] as

\[
L(\theta) = \prod_{i=1}^{k} f(x_i; \theta).
\]

(3)

Let \( X_{U(1)}, \ldots, X_{U(k)} \) be the first \( k \) upper record values from the two-parameter Rayleigh distribution. The likelihood function based on record values is given by

\[
L(\mu, \sigma) = \exp \left[ -\frac{(X_{U(k)} - \mu)^2}{2\sigma^2} \right] \prod_{i=1}^{k} \frac{x_{U(i)} - \mu}{\sigma^2}.
\]

(4)

The MLEs \( \hat{\mu} \) and \( \hat{\sigma} \) can be found by solving the following likelihood equations for \( \mu \) and \( \sigma \) simultaneously:

\[
\frac{\partial}{\partial \mu} \log L(\mu, \sigma) = \frac{x_{U(k)} - \mu}{\sigma^2} - \frac{1}{2} \sum_{i=1}^{k} \frac{1}{x_{U(i)} - \mu} = 0,
\]

(5)

\[
\frac{\partial}{\partial \sigma} \log L(\mu, \sigma) = \frac{(x_{U(k)} - \mu)^2}{\sigma^3} - \frac{2k}{\sigma} = 0.
\]

(6)

On the other hand, the MLEs cannot be expressed in closed form and their exact distributions are difficult to derive. Alternatively, by the asymptotic normality of the MLE, the approximate 100(1 - \(\alpha\))% CIs for \( \mu \) and \( \sigma \) can be obtained as

\[
\hat{\mu} \pm Z_{\alpha/2} \sqrt{\text{Var}(\hat{\mu})},
\]

\[
\hat{\sigma} \pm Z_{\alpha/2} \sqrt{\text{Var}(\hat{\sigma})},
\]

where \( Z_{\alpha/2} \) denotes the upper \( \alpha/2 \) point of the standard normal distribution and the variances \( \text{Var}(\hat{\mu}) \) and \( \text{Var}(\hat{\sigma}) \) are the diagonal elements of the asymptotic variance-covariance matrix obtained by inverting the Fisher information matrix for unknown parameters \((\mu, \sigma)\):

\[
I(\mu, \sigma) = \begin{bmatrix}
E\left(-\frac{\partial^2}{\partial \mu^2} \log L(\mu, \sigma)\right) & E\left(-\frac{\partial^2}{\partial \mu \partial \sigma} \log L(\mu, \sigma)\right) \\
E\left(-\frac{\partial^2}{\partial \sigma \partial \mu} \log L(\mu, \sigma)\right) & E\left(-\frac{\partial^2}{\partial \sigma^2} \log L(\mu, \sigma)\right)
\end{bmatrix},
\]

(7)

under certain regularity conditions. Nevertheless, it can provide inappropriate results because the supports of \( \mu \) and \( \sigma \) do not coincide with that of the normal distribution and the record values are rarely observed, as mentioned before.

### 2.1. Confidence Interval

This subsection develops inference methods based on the pivotal quantities to construct exact CIs for unknown parameters \((\mu, \sigma)\) and PIs for the future upper record values. Note that the proposed method is much easier to calculate than the maximum likelihood method. The following provides some pivotal quantities.

Let

\[
Z_i = -\log \left[1 - F(x_{U(i)})\right] = \frac{(X_{U(i)} - \mu)^2}{2\sigma^2},
\]

(8)

\( i = 1, \ldots, k \).
\( Z_1 < \cdots < Z_k \) are the upper record values with a standard exponential distribution. From this result, the following spacing can be obtained:

\[
S_i = Z_i - Z_{i-1} = \frac{1}{2\sigma^2} \left[ (X_{U(i)} - \mu)^2 - (X_{U(i-1)} - \mu)^2 \right], \quad i = 1, \ldots, k \quad (Z_0 \equiv 0),
\]

which are the iid RVs from the standard exponential distribution (see Arnold et al., [10]). Based on the spacing, a pivotal quantity can be obtained:

\[
V = 2^{k-1} \sum_{i=2}^{k} S_i = \frac{(X_{U(k)} - \mu)^2}{\sigma^2} - \frac{(X_{U(1)} - \mu)^2}{\sigma^2},
\]

having the \( \chi^2 \) distribution with \( 2(k-1) \) degrees of freedom. Because they have independent RVs, the following pivotal quantity is obtained:

\[
W(\mu) = \frac{V/2(k-1)}{T/2} = \frac{1}{k-1} \left[ \frac{(X_{U(k)} - \mu)^2}{X_{U(1)} - \mu} - 1 \right],
\]

which has a \( F \) distribution with \( 2(k-1) \) and 2 degrees of freedom. An exact \( 100(1-\alpha)\% \) CI for \( \mu \) based on the pivotal quantity \( W(\mu) \) can be constructed as

\[
\left( \frac{X_{U(k)} - X_{U(1)} \sqrt{(k-1)F_{\alpha/2,(2(k-1),2)} + 1}}{1 - \sqrt{(k-1)F_{\alpha/2,(2(k-1),2)} + 1}} \right) < \mu < \left( \frac{X_{U(k)} - X_{U(1)} \sqrt{(k-1)F_{1-\alpha/2,(2(k-1),2)} + 1}}{1 - \sqrt{(k-1)F_{1-\alpha/2,(2(k-1),2)} + 1}} \right),
\]

for any \( 0 < \alpha < 1 \), where \( F_{\alpha/2,(2(k-1),2)} \) is the upper \( \alpha \) percentile of the \( F \) distribution with \( \nu_1 \) and \( \nu_2 \) degrees of freedom.

Moreover, because \( Q(\sigma) = T + V \) has the \( \chi^2 \) distribution with \( 2k \) degrees of freedom, an exact 100\((1-\alpha)\)% CI for \( \sigma \) based on the pivotal quantity \( Q(\sigma) \) can be constructed as

\[
\left( \frac{(X_{U(k)} - \mu)^2}{\chi^2_{\alpha, k}} \right)^{1/2} < \sigma < \left( \frac{(X_{U(k)} - \mu)^2}{\chi^2_{1-\alpha, k}} \right)^{1/2},
\]

for any \( 0 < \alpha < 1 \),

where \( \chi^2_{\alpha, k} \) is the upper \( \alpha \) percentile of the \( \chi^2 \) distribution with \( k \) degrees of freedom. Note that because the precise CI (13) depends on the nuisance parameter \( \mu \), this paper shows how to address the nuisance parameter \( \mu \) based on a generalized pivotal quantity, and an exact CI for \( \sigma \) is proposed based on the generalized pivotal quantity.

Let \( \mu^* \) be the unique solution of \( W(\mu) = W \), where \( W \) has a \( F \) distribution with \( 2(k-1) \) and 2 degrees of freedom. The unique solution can then be given by

\[
\mu^* = \frac{X_{U(k)} - X_{U(1)} \sqrt{(k-1)W + 1}}{1 - \sqrt{(k-1)W + 1}}.
\]

Moreover, let \( Q \) be the RV from the \( \chi^2 \) distribution with \( 2k \) degrees of freedom. The generalized pivotal quantity from the pivotal quantity \( Q(\sigma) \) is given by

\[
W(\mu^*) = \frac{(X_{U(k)} - \mu^*)^2}{Q}.
\]

Here, the samples \( W(\mu^*_1), \ldots, W(\mu^*_N) \) can be obtained by generating \( N(\geq 10,000) \) the RVs \( W \) and \( Q \). \( W(\mu^*_1), \ldots, W(\mu^*_N) \) are ordered as \( W(\mu^*_1), \ldots, W(\mu^*_N) \). Therefore, an exact \( 100(1-\alpha)\% \) CI for \( \sigma \) based on the generalized pivotal quantity \( W(\mu^*) \) can be constructed:

\[
\left( \frac{W(\mu^*_1)}{[N/(100)] \times \alpha / 2}, \frac{W(\mu^*_N)}{[N/(100)] \times (1-\alpha / 2)} \right),
\]

where \( [z] \) denotes the largest integer less than or equal to \( z \).

In Section 3, the proposed CIs are examined in terms of the coverage probability (CPS) to determine if they are valid CIs.

2.2. Predictive Interval. This subsection develops a method for predicting the future upper record values based on the observed upper record values \( x_{U(1)}, \ldots, x_{U(k)} \) by providing a pivotal quantity. Let \( X_{U(s)} \) be a future upper record value. The conditional density function of \( X_{U(s)} \), given \( x_{U(k)} \), defined by Ahsanullah [11], is given by

\[
f_{X_{U(s)}|x_{U(k)}}(x_{U(s)} \mid \mu, \sigma) = \frac{1}{\Gamma(s-k)} \left[ \log (1 - F(x_{U(k)})) \right]
- \log (1 - F(x_{U(s)}))^s - \frac{(s-k-1)}{1 - F(x_{U(k)})}.
\]

from the Markov property of the record values. Assuming that the observed upper record values, \( x_{U(1)}, \ldots, x_{U(k)} \), arise
from the Rayleigh distribution with the pdf (2), the conditional density function (17) is written as
\[ f_{X_U(s) | X_U(k)}(X_U(s) | \mu, \sigma) = \frac{1}{\Gamma(s-k)} \left( \frac{(x_{U(s)} - \mu)^2}{2\sigma^2} - \frac{(x_{U(k)} - \mu)^2}{2\sigma^2} \right)^{s-k-1} \exp \left( \frac{(x_{U(s)} - \mu)^2}{2\sigma^2} - \frac{(x_{U(k)} - \mu)^2}{2\sigma^2} \right). \]

Let
\[ Y = \frac{1}{\sigma^2} \left( (X_{U(s)} - \mu)^2 - (X_{U(k)} - \mu)^2 \right). \]

Because the Jacobian of transformation is
\[ \frac{d}{dy} X_{U(s)} = \frac{\sigma^2}{2\sqrt{\sigma^2 y + (x_{U(k)} - \mu)^2}}, \]
the density function of \( Y \) is given by
\[ f(y) = \frac{1}{2^{s-k-1} \Gamma(s-k)} y^{s-k-1} e^{-y/2}, \quad 0 < y < \infty, \]
which is the pdf of the \( \chi^2 \) distribution with \( 2(s-k) \) degrees of freedom. Suppose that \( \mu \) and \( \sigma \) are known. An exact 100(1 - \( \alpha \))% CI based on the pivotal quantity \( Y \) for the future upper record value \( X_{U(s)} \) is obtained as
\[ \left( \mu + \sqrt{\sigma^2 \chi^2_{1-\alpha/2(s-k)}} + (x_{U(k)} - \mu)^2, \mu + \sqrt{\sigma^2 \chi^2_{1-\alpha/2(s-k)}} + (x_{U(s)} - \mu)^2 \right). \]

When \( \mu \) and \( \sigma \) are unknown, they can be substituted with \( \mu^* \) and \( W(\mu^*) \) in PI (22) based on the fact that \( W(\mu^*) \) is the generalized pivotal quantity for constructing the exact CI for \( \sigma \). In the same way, the generalized pivotal quantity is given by
\[ Y(\mu^*) = \left( \mu^* + \sqrt{W(\mu^*)^2 \chi^2_{1-\alpha/2(s-k)}} + (x_{U(k)} - \mu^*)^2 \right), \]
and an exact 100(1 - \( \alpha \))% PI for \( X_{U(s)} \) based on the generalized pivotal quantity \( Y(\mu^*) \) can be constructed as follows:
\[ \left( Y(\mu^*)_{(N/100)(1-\alpha/2)}, Y(\mu^*)_{(N/100)(1-\alpha/2)} \right). \]

3. Application

This section assesses the proposed methods through a Monte Carlo simulation and presents a real data set.

3.1. Simulation Study. The proposed exact CIs (12) and (16) are assessed in terms of their CPs and average lengths (ALs). The upper record values were first generated from the standard Rayleigh distribution with \( \mu = 0 \) and \( \sigma = 1 \) for different \( k \), and the CIs (12) and (16) were calculated based on the generated samples by using the provided methods in Section 2.1. The CPs and ALs of the exact CIs were obtained over 10,000 simulations. These values are reported in Table 1. Table 1 shows that the CPs matched their corresponding nominal levels even in a small sample size and that all ALs decrease with increasing sample size.

3.2. Real Data. To illustrate the proposed inference procedure, the survival times in (days) of a group of lung cancer patients (from Lawless [12, p. 319]) were considered as follows:

\[ 6.96, 9.30, 6.96, 7.24, 9.30, 4.90, 8.42, 6.05, 10.18, 6.82, 8.58, 7.77, 11.94, 11.25, 12.94, \text{ and } 12.94. \]

From the data, the observed upper record values were 6.96, 9.30, 10.18, 11.94, and 12.94. Soliman and Al-Aboud [13] showed that the Rayleigh distribution fits the observed record values well. These record values are employed to obtain the proposed CIs (12) and (16). Moreover, the exact PIs for the future upper record values \( X_{U(s)} \) (\( s = 6, 7 \)) were computed, as listed in Table 2.

4. Concluding Remarks

This paper proposes methods for inferencing the exact CIs for unknown parameters \( (\mu, \sigma) \) in the Rayleigh distribution based on the upper record values and exact PIs for the future upper record values by providing some pivotal quantities.

Because the proposed exact CI (13) and PI (22) depend on the nuisance parameters, this study proposed generalized pivotal quantities \( W(\mu^*) \) and \( Y(\mu^*) \) to solve the drawback. The proposed methods were more computationally convenient than the maximum likelihood method. Moreover, the proposed exact CIs provide very good performance even in small sample sizes. If the location parameter of the Rayleigh distribution is of interest, the exact CI (12) should be used because it does not have any nuisance parameter.

Competing Interests

The authors declare that they have no competing interests.
Table 1: CPs (ALs) of exact 95% CIs for $\mu$ and $\sigma$.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\mu$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.949 (6.135)</td>
<td>0.952 (2.468)</td>
</tr>
<tr>
<td>7</td>
<td>0.953 (4.926)</td>
<td>0.954 (1.723)</td>
</tr>
<tr>
<td>9</td>
<td>0.951 (4.353)</td>
<td>0.950 (1.329)</td>
</tr>
<tr>
<td>11</td>
<td>0.947 (4.044)</td>
<td>0.953 (1.127)</td>
</tr>
<tr>
<td>13</td>
<td>0.949 (3.845)</td>
<td>0.954 (0.995)</td>
</tr>
<tr>
<td>15</td>
<td>0.950 (3.688)</td>
<td>0.951 (0.901)</td>
</tr>
</tbody>
</table>

Table 2: Results for real data.

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$X_{(16)}$</th>
<th>$X_{(27)}$</th>
</tr>
</thead>
</table>

References


Submit your manuscripts at http://www.hindawi.com