JIRSS (2017)
Vol. 16, No. 2, pp 1-19
DOI:10.22034/jirss.2017.16.01

Prediction for Lindley Distribution Based on Type-II Right Censored Samples

Reza Valiollahi [1], **Akbar Asgharzadeh** [2], **Hon K. T. Ng** [3]

[1]Department of Statistics, Semnan University, Semnan, Iran
[2]Department of Statistics, University of Mazandaran, Babolsar, Iran
[3]Department of Statistical Science, Southern Methodist University, Dallas, USA
Received: 06/06/2016, Revision received: 19/03/2017, Published online: 27/11/2017

Abstract. Lindley distribution has received considerable attention in the statistical literature due to its simplicity. In this paper, we consider the problem of predicting the failure times of experimental units that are censored in a right-censored sample when the underlying lifetime is Lindley distributed. The maximum likelihood predictor, the Best unbiased predictor and the conditional median predictor are derived. Prediction intervals based on these predictors are considered. We further propose two resampling-based procedures for obtaining the prediction intervals. A numerical example is used to illustrate the methodology developed in this paper. Finally, a Monte Carlo simulation study is employed to evaluate the performance of different prediction methods.

Keywords. Best unbiased predictor, Conditional median prediction, Highest conditional density, Maximum likelihood prediction, Prediction interval.

MSC: 62G30, 62F10, 62M20.

Corresponding Author: A. Asgharzadeh (a.asgharzadeh@umz.ac.ir)
R. Valiollahi (rvaliollahi@semnan.ac.ir)
H. K. T. Ng (ngh@mail.smu.edu)

1 Introduction

The Lindley distribution was first proposed by Lindley (1958) in the context of Bayesian inference. A random variable having a Lindley distribution has the probability density function (pdf)

$$f(x;\theta) = \frac{\theta^2}{1+\theta}(1+x)e^{-\theta x}, \quad x > 0, \ \theta > 0, \tag{1.1}$$

and the cumulative distribution function (cdf)

$$F(x;\theta) = 1 - \left(\frac{1+\theta+\theta x}{1+\theta}\right)e^{-\theta x}, \quad x > 0, \ \theta > 0. \tag{1.2}$$

A distribution that is close in from to the Lindley distribution is the well-known exponential distribution. Ghitany *et al.* (2008) discussed different properties of the Lindley distribution and showed that in many ways this distribution is better than the exponential distribution in applications. In recent years, the Lindley distribution has been studied and generalized by many authors; see, for example, Ghitany *et al.* (2008), Zakerzadeh and Dolati (2009), Ghitany *et al.* (2011), Nadarajah *et al.* (2011), Bakouch *et al.* (2012), Gupta and Singh (2013), Gomez-Deniz *et al.* (2014) and Asgharzadeh *et al.* (2016a).

In recent years, the parameter estimation problem for the Lindley distribution has been discussed by several authors. Krishna and Kumar (2011) discussed the estimation of the parameter of Lindley distribution under progressive Type-II censoring. Al-Mutairi *et al.* (2013) discussed the estimation of the stress-strength parameter $R = \Pr(Y < X)$, when both stress and strength variables follow the Lindley distribution. Gupta and Singh (2013) considered the maximum likelihood and Bayes estimators for the Lindley parameter based on hybrid censored data. Ali *et al.* (2013) considered the Bayesian estimation for the Lindley model with informative and noninformative priors under different loss functions. Recently, Asgharzadeh *et al.* (2016b) discussed the maximum likelihood and Bayesian estimation for the Lindley model based on record data.

Although the estimation of the parameter of the Lindley distribution has been discussed extensively in the literature, to the best of our knowledge, the prediction problem for future failures has not been considered before. Specifically, suppose that n items are placed on a life-testing experiment and it is planned that the experiment will be terminated as soon as the m-th (where m is pre-fixed) failure is observed. Then, only the first m failures out of n units under the test will be observed. The data obtained from such a life-test, denoted as $X_{1:n} < X_{2:mn} < \ldots < X_{m:n}$, is referred to as a *Type-II censored sample*. We are aimed in predicting the future failures $Y = X_{s+m:n}(s = 1, 2, \ldots, n-m)$ based on the Type-II censored sample $\mathbf{X} = (X_{1:n}, X_{2:n}, \ldots, X_{m:n})$. We obtain the maximum likelihood predictor, the best unbiased predictor and the conditional median predictor. Two prediction intervals are also proposed. We further propose two resampling-based procedures for obtaining the point and interval predictors.

The paper is organized as follows. Section 2 provides some preliminaries. In Section 3, we derive the maximum likelihood predictor. In Section 4, the conditional approach is used to obtain the best unbiased predictor and the conditional median predictor. In Section 5, we propose two procedures for obtaining the prediction intervals. In Section 6, the resampling-based procedures are proposed for obtaining the point and interval predictors. In Section 7, one real data analysis has been presented for illustrative purposes. Finally, in Section 8, Monte Carlo simulations are performed to compare the performances of the different proposed predictors.

2 Some Preliminaries

Let $\mathbf{X} = (X_{1:n}, X_{2:n}, \ldots, X_{m:n})$ denote a Type-II right censored sample from model (1.1). For notational simplicity, we denote the ordered Type-II censored sample as $(X_1, X_2, \ldots, X_m)$ by suppressing the order statistic notation. Based on the observed data, the likelihood function for θ can be expressed as

$$\begin{aligned} L(\theta) &\propto \prod_{i=1}^{m}\left[\frac{\theta^2}{1+\theta}(1+x_i)e^{-\theta x_i}\right]\left(\frac{1+\theta+\theta x_m}{1+\theta}e^{-\theta x_m}\right)^{n-m} \\ &= \frac{\theta^{2m}}{(1+\theta)^n}\prod_{i=1}^{m}(1+x_i)(1+\theta+\theta x_m)^{n-m}e^{-\theta\left[\sum_{i=1}^{m} x_i+(n-m)x_m\right]}. \end{aligned} \tag{2.3}$$

The log-likelihood function is

$$\begin{aligned} \ln L(\theta) &= \text{constant} + 2m\ln\theta - n\ln(1+\theta) + (n-m)\ln(1+\theta+\theta x_m) \\ &- \theta\left[\sum_{i=1}^{m} x_i + (n-m)x_m\right]. \end{aligned} \tag{2.4}$$

Then, we can obtain the log-likelihood equation as

$$\frac{d\ln L(\theta)}{d\theta} = \frac{2m}{\theta} - \frac{n}{1+\theta} + \frac{(n-m)(1+x_m)}{1+\theta+\theta x_m} - \left[\sum_{i=1}^{m} x_i + (n-m)x_m\right] = 0. \tag{2.5}$$

We can see that the maximum likelihood estimator (MLE) of θ, $\widehat{\theta}_{MLE}$, cannot be obtained in a closed form. Therefore, the MLE of θ needs to be obtained using numerical methods.

Instead of estimating the parameter in the Lindley distribution, the main objective here is to predict the future life-lengths $Y = X_{s+m:n}(s = 1, 2, \ldots, n-m)$ based on $\mathbf{X} = (X_1, ..., X_m)$. As we shall see later, the conditional distribution of Y given $\mathbf{X} = \mathbf{x}$ plays an important role to predict Y. By the Markovian property of Type-II right censored order statistics, the conditional distribution of Y given $\mathbf{X} = \mathbf{x}$ is the same as the conditional distribution of Y given $X_m = x_m$. This implies that the conditional pdf of Y given $\mathbf{X} = \mathbf{x}$ is the same as the pdf of the s-th order statistic from a sample of size $n-m$ from the population with pdf $f(y;\theta)/[1-F(x_m;\theta)]$, $y \geq x_m$ (i.e., a left-truncated pdf truncated at x_m). Therefore, the conditional pdf of Y given $\mathbf{X} = \mathbf{x}$ can be written as

$$\begin{aligned} g(y|\mathbf{x};\theta) &= s\binom{n-m}{s} f(y)\left[F(y)-F(x_m)\right]^{s-1} \\ &\quad \times \left[1-F(y)\right]^{n-m-s}\left[1-F(x_m)\right]^{-(n-m)}, \quad y \geq x_m. \end{aligned} \tag{2.6}$$

For the Lindely distribution, Equation (2.6) becomes

$$\begin{aligned} g(y|\mathbf{x};\theta) &= s\binom{n-m}{s}\theta^2(1+y)(1+\theta+\theta y)^{n-m-s} \\ &\quad \times e^{-\theta(n-m-s+1)y}e^{\theta(n-m)x_m}(1+\theta+\theta x_m)^{-(n-m)} \\ &\quad \times\left[(1+\theta+\theta x_m)e^{-\theta x_m}-(1+\theta+\theta y)e^{-\theta y}\right]^{s-1}. \end{aligned} \tag{2.7}$$

3 Maximum Likelihood Approach

In this section, we derive the maximum likelihood predictor (MLP) $Y = X_{s+m:n}(s = 1, 2, \ldots, n-m)$, on the basis of the Type-II censored sample $\mathbf{X} = (X_1, ..., X_m)$ via likelihood approach. The likelihood approach was first proposed by Kaminsky and Rhodin (1985) to predict the future order statistics and estimate the parameters involved in the model. Given the informative sample $\mathbf{x}$, the predictive likelihood function (PLF) of Y and θ is considered and maximized simultaneously with regard to future observation Y and the parameter θ. The PLF of Y and θ, is given by

$$L(y,\theta|\mathbf{x}) = g(y|\mathbf{x};\theta)f(\mathbf{x};\theta). \tag{3.8}$$

Suppose $\widehat{Y} = u(\mathbf{X})$ and $\widehat{\theta} = v(\mathbf{X})$ are statistics for which

$$L(u(\mathbf{x}), v(\mathbf{x})|\mathbf{x}) = \sup_{(y,\theta)} L(y, \theta|\mathbf{x}). \tag{3.9}$$

Then, we call $u(\mathbf{X})$ the MLP of Y and $v(\mathbf{X})$ the predictive maximum likelihood estimator (PMLE) of θ.

For the Lindley model, by substituting Equations (2.3) and (2.7) into (3.8), the PLF of Y and θ can be obtained as

$$\begin{aligned} L(y, \theta|\mathbf{x}) \propto\ & s\binom{n-m}{s}\frac{\theta^{2m+2}}{(1+\theta)^n} \\ & \times \prod_{i=1}^{m}(1+x_i)(1+y)e^{-\theta\{(n-m-s+1)y+\sum_{i=1}^{m}x_i\}} \\ & \times(1+\theta+\theta y)^{n-m-s} \\ & \times\left[(1+\theta+\theta x_m)e^{-\theta x_m} - (1+\theta+\theta y)e^{-\theta y}\right]^{s-1}. \end{aligned} \tag{3.10}$$

The predictive log-likelihood function is given by

$$\begin{aligned} \ln L(y, \theta|\mathbf{x}) =\ & \text{constant} + (2m+2)\ln\theta - n\ln(1+\theta) \\ & + \ln(1+y) + (n-m-s)\ln(1+\theta+\theta y) \\ & + (s-1)\ln\left[(1+\theta+\theta x_m)e^{-\theta x_m} - (1+\theta+\theta y)e^{-\theta y}\right] \\ & - \theta\left[(n-m-s+1)y + \sum_{i=1}^{m} x_i\right]. \end{aligned} \tag{3.11}$$

The predictive likelihood equations (PLEs) are obtained by differentiating $\ln L(y, \theta|\mathbf{x})$ in (3.11) with respect to y and θ, and they are as follows:

$$\begin{aligned} \frac{\partial \ln L(y, \theta|\mathbf{x})}{\partial y} =\ & \frac{1}{y+1} - \theta(n-m-s+1) + \frac{(n-m-s)\theta}{1+\theta+\theta y} \\ & + (s-1)\frac{\theta^2(1+y)e^{-\theta y}}{(1+\theta+\theta x_m)e^{-\theta x_m} - (1+\theta+\theta y)e^{-\theta y}} \\ =\ & 0, \end{aligned} \tag{3.12}$$

$$\begin{aligned}\frac{\partial \ln L(y,\theta|\mathbf{x})}{\partial\theta} &= \frac{(s-1)e^{-\theta x_m}[1-\theta x_m(1+x_m)]-e^{-\theta y}[1-\theta y(1+y)]}{(1+\theta+\theta x_m)e^{-\theta x_m}-(1+\theta+\theta y)e^{-\theta y}}\\ &+ \frac{2m+2}{\theta}-\frac{n}{1+\theta}-(n-m-s+1)y-\sum_{i=1}^{m}x_i\\ &+ (n-m-s)\frac{1+y}{1+\theta+\theta y}\\ &= 0. \end{aligned} \tag{3.13}$$

By solving Equations (3.12) and (3.13) with respect to y and θ simultaneously, the MLP of Y, $\widehat{Y}_{MLP}$, and PMLE of θ can be obtained. Numerical methods can be used to solve these PLEs and obtain the MLP $\widehat{Y}_{MLP}$ and PMLE of θ.

4 Conditional Distribution Approach

Here, the conditional distribution of Y given $\mathbf{X}=\mathbf{x}$ is used to predict Y. A statistic $\widehat{Y}$ which is used to predict $Y = X_{s+m:n}$ is called a best unbiased predictor (BUP) of Y if the prediction error $(\widehat{Y}-Y)$ has a zero mean and the prediction variance, $\mathrm{Var}(\widehat{Y}-Y)$, is smaller than or equal to that of any other unbiased predictor of Y. The BUP of Y is the mean of the conditional distribution of Y given $\mathbf{X}=\mathbf{x}$. Therefore, the BUP is given by

$$\widehat{Y}_{BUP} = E(Y|\mathbf{X}) = \int_{x_m}^{\infty} y g(y|\mathbf{x};\theta)dy. \tag{4.14}$$

Using (2.7) and the binomial expansion

$$\begin{aligned} &\left[(1+\theta+\theta x_m)e^{-\theta x_m}-(1+\theta+\theta y)e^{-\theta y}\right]^{s-1}\\ =& \sum_{j=0}^{s-1}\left[\binom{s-1}{j}(-1)^{(s-j-1)}(1+\theta+\theta x_m)^j\right.\\ &\left.\times(1+\theta+\theta y)^{s-j-1}e^{-\theta j x_m}e^{-\theta(s-j-1)y}\right], \end{aligned} \tag{4.15}$$

we can obtain the BUP of Y as

$$\begin{aligned}\widehat{Y}_{BUP} &= s\binom{n-m}{s}\theta^2 \\ &\quad \times \sum_{j=0}^{s-1}\binom{s-1}{j}(-1)^{(s-j-1)}(1+\theta+\theta x_m)^{-(n-m-j)}e^{\theta(n-m-j)x_m} \\ &\quad \times \int_{x_m}^{\infty} y(1+y)(1+\theta+\theta y)^{n-m-j-1}e^{-\theta(n-m-j)y}dy. \end{aligned} \tag{4.16}$$

Once again, by using the binomial expansion, the last integral can be written as

$$\begin{aligned}&\int_{x_m}^{\infty} y(1+y)(1+\theta+\theta y)^{n-m-j-1}e^{-\theta(n-m-j)y}dy \\ &= \sum_{l=0}^{n-m-j-1}\binom{n-m-j-1}{l}\theta^l(1+\theta)^{n-m-j-l-1} \\ &\quad \times\left[\frac{\Gamma(l+1,\theta(n-m-j)x_m)}{\theta^{l+1}(n-m-j)^{l+1}}+\frac{\Gamma(l+2,\theta(n-m-j)x_m)}{\theta^{l+2}(n-m-j)^{l+2}}\right], \end{aligned} \tag{4.17}$$

where $\Gamma(s,x)$ is the incomplete function defined as

$$\Gamma(s,x)=\int_{x_m}^{\infty} t^{s-1}e^{-t}dt.$$

Replacing (4.17) in (4.16), the BUP is obtained as

$$\begin{aligned}\widehat{Y}_{BUP} &= s\binom{n-m}{s}\theta^2\sum_{j=0}^{s-1}\sum_{l=0}^{n-m-j-1}\binom{s-1}{j}\binom{n-m-j-1}{l}(-1)^{(s-j-1)} \\ &\times \quad \theta^l(1+\theta)^{n-m-j-l-1}(1+\theta+\theta x_m)^{-(n-m-j)}e^{\theta(n-m-j)x_m} \\ &\times \quad \left[\frac{\Gamma(l+1,\theta(n-m-j)x_m)}{\theta^{l+1}(n-m-j)^{l+1}}+\frac{\Gamma(l+2,\theta(n-m-j)x_m)}{\theta^{l+2}(n-m-j)^{l+2}}\right]. \end{aligned} \tag{4.18}$$

If the parameter θ is unknown, the BUP of Y can be approximated by replacing θ by its MLE.

Another conditional predictor is the conditional median predictor (CMP). This predictor was first proposed by Raqab and Nagaraja (1995) in the context of order statistics. We call a predictor $\widehat{Y}$ is a CMP of Y if it is the median of the conditional distribution of Y given $\mathbf{X} = \mathbf{x}$, i.e.,

$$\Pr_\theta(Y \le \widehat{Y} | \mathbf{X} = \mathbf{x}) = \Pr_\theta(Y \ge \widehat{Y} | \mathbf{X} = \mathbf{x}). \tag{4.19}$$

For Lindley distribution, we can obtain

$$\Pr_\theta(Y \le \widehat{Y} | \mathbf{X} = \mathbf{x}) = \\ \Pr_\theta\left(1 - \frac{(1+\theta+\theta Y)e^{-\theta Y}}{(1+\theta+\theta X_m)e^{-\theta X_m}} \le 1 - \frac{(1+\theta+\theta\widehat{Y})e^{-\theta\widehat{Y}}}{(1+\theta+\theta X_m)e^{-\theta X_m}} \middle| \mathbf{X} = \mathbf{x}\right). \tag{4.20}$$

Note that the conditional distribution

$$\frac{F(Y) - F(x_m)}{1 - F(x_m)} = 1 - \frac{(1+\theta+\theta Y)e^{-\theta Y}}{(1+\theta+\theta x_m)e^{-\theta x_m}},$$

given $\mathbf{X} = \mathbf{x}$, is a beta distribution with parameters s and $n - m - s + 1$ (denoted as $Beta(s, n - m - s + 1)$) using an application of probability integral transformation and the fact that the ith smallest order statistics from a random sample of size n from the standard uniform distribution is distributed as *Beta(i,n-i+1)*. Therefore, by (4.20), the CMP of Y can be obtained by solving the equation

$$1 - \frac{(1+\theta+\theta\widehat{Y})e^{-\theta\widehat{Y}}}{(1+\theta+\theta x_m)e^{-\theta x_m}} = Med(B), \tag{4.21}$$

where B is a random variable which follows a $Beta(s, n - m - s + 1)$ distribution and $Med(B)$ stands for the median of B. Now, from (4.21), the CMP of Y, $\widehat{Y}_{CMP}$, is

$$\widehat{Y}_{CMP} = -1 - \frac{1}{\theta} - \frac{1}{\theta}\, W_{-1}\left([Med(B) - 1](1+\theta+\theta x_m)e^{-(1+\theta+\theta x_m)}\right), \tag{4.22}$$

where $W_{-1}(\cdot)$ denotes the negative branch of the Lambert W function, i.e., the solution of the equation $W(z)e^{W(z)} = z$. For more details on the Lambert W function, see Chapeau-Blondeau and Monir (2002). When θ is unknown, we can substitute θ with its MLE and obtain an approximation of the CMP of Y.

5 Prediction Intervals

Given the Type-II censored sample $\mathbf{X} = (X_1, ..., X_m)$, we want to obtain the prediction intervals for the sth censored value $Y = X_{s+m:n}(s = 1, 2, \cdots, n-m)$. A $(1-\alpha)100\%$ predictive interval for Y is an interval $(L(\mathbf{X}), U(\mathbf{X}))$ such that $\Pr(L(\mathbf{X}) < Y < U(\mathbf{X})) = 1-\alpha$. In this section, two types of prediction intervals are constructed. We first provide a pivotal quantity. Consider the random variable Z as

$$Z = 1 - \frac{(1+\theta+\theta Y)e^{-\theta Y}}{(1+\theta+\theta X_m)e^{-\theta X_m}}. \tag{5.23}$$

As discussed in Section 4, the distribution of Z given $\mathbf{X} = \mathbf{x}$ is a $Beta(s, n-m-s+1)$ distribution. Therefore, we can consider Z as a pivotal quantity for the construction of the prediction interval for Y. From the pivotal quantity Z, it follows that

$$\Pr(B_{\frac{\alpha}{2}} < Z < B_{1-\frac{\alpha}{2}} | \mathbf{X}) = 1-\alpha,$$

where B_α is the 100α-th percentile of $Beta(s, n-m-s+1)$ distribution. By solving the inequalities for Y, an exact $(1-\alpha)100\%$ predictive interval for Y is $(L_1(\mathbf{X}), U_1(\mathbf{X}))$ where

$$L_1(\mathbf{X}) = -1 - \frac{1}{\theta} - \frac{1}{\theta} W_{-1}\left([B_{\frac{\alpha}{2}} - 1](1+\theta+\theta x_m)e^{-(1+\theta+\theta x_m)}\right), \tag{5.24}$$

$$U_1(\mathbf{X}) = -1 - \frac{1}{\theta} - \frac{1}{\theta} W_{-1}\left([B_{1-\frac{\alpha}{2}} - 1](1+\theta+\theta x_m)e^{-(1+\theta+\theta x_m)}\right). \tag{5.25}$$

Since θ is unknown, the prediction limits $L_1(\mathbf{X})$ and $U_1(\mathbf{X})$ can be approximated by replacing θ with its corresponding MLE.

Let us now consider the highest conditional density (HCD) method for constructing the prediction interval of Y. The distribution of Z given $\mathbf{X} = \mathbf{x}$ is a $Beta(s, n-m-s+1)$ distribution with pdf

$$g(z) = \frac{z^{s-1}(1-z)^{n-m-s}}{Beta(s, n-m-s+1)}, \quad 0 < z < 1,$$

which is a unimodal function of z, for $1 < s < n-m$. A $(1-\alpha)100\%$ HCD predictive interval for Y is $(L_2(\mathbf{X}), U_2(\mathbf{X}))$, where $L_2(\mathbf{X})$ and $U_2(\mathbf{X})$ are given by

$$L_2(\mathbf{X}) = -1 - \frac{1}{\theta} - \frac{1}{\theta} W_{-1}\left([d_1 - 1](1+\theta+\theta X_m)e^{-(1+\theta+\theta X_m)}\right), \tag{5.26}$$

$$U_2(\mathbf{X}) = -1 - \frac{1}{\theta} - \frac{1}{\theta} W_{-1}\left([d_2 - 1](1+\theta+\theta X_m)e^{-(1+\theta+\theta X_m)}\right), \tag{5.27}$$

and d_1 and d_2 satisfy the equations

$$\int_{d_1}^{d_2} g(z)dz = 1 - \alpha \tag{5.28}$$

and

$$g(d_1) = g(d_2). \tag{5.29}$$

The Equations (5.28) and (5.29) can be simplified as

$$B_{d_2}(s, n-m-s+1) - B_{d_1}(s, n-m-s+1) = 1 - \alpha \tag{5.30}$$

and

$$\left(\frac{1-d_2}{1-d_1}\right)^{n-m-s} = \left(\frac{d_1}{d_2}\right)^{s-1}, \tag{5.31}$$

where

$$B_t(a,b) = \frac{1}{B(a,b)} \int_0^t x^{a-1}(1-x)^{b-1}\, dx$$

is the incomplete beta function. Note that, for the case $s = 1$ or $s = n - m$, the function $g(z)$ is not unimodal and the HCD prediction interval cannot be obtained.

6 Resampling Approach

In this section, we construct both point and interval predictions of future order statistics based on the percentile parametric bootstrap method (Efron and Tibshirani , 1993). In addition, we develop a bias-adjusted estimator based on the parametric bootstrap. To obtain the percentile bootstrap confidence intervals, we use the following algorithm:

1. Based on the original sample $\mathbf{x} = (x_1, x_2, \ldots, x_m)$, obtain $\hat{\theta}$, the MLE of θ.

2. Generate a random sample of size $n - m$, denoted as $\mathbf{y} = (y_1, y_2, \ldots, y_{n-m})$ from a left-truncated distribution with the point of truncation to be x_m, i.e., with the p.d.f.

$$f_{x_m}(x) = \frac{f(x)}{1 - F(x_m)}, x > x_m.$$

3. Obtain $\mathbf{y}^{(1)} = (x^{(1)}_{m+1:n}, x^{(1)}_{m+2:n}, \ldots, x^{(1)}_{n:n})$ from Lindley distribution with parameter $\hat{\theta}$ by the inverse transformation method: $x^{(1)}_{s:n} = F^{-1}(u_{s:n})$.

4. Repeat Steps 2 – 3 for B times and obtain ${x_{s:n}}^{(1)}, {x_{s:n}}^{(2)}, \ldots, {x_{s:n}}^{(B)}$, for $s = m+1, m+2, \ldots, n$.
5. Arrange ${x_{s:n}}^{(1)}$, ${x_{s:n}}^{(2)}$, …, ${x_{s:n}}^{(B)}$ in ascending order and obtain ${x_{s:n}}^{[1]}$, ${x_{s:n}}^{[2]}$, …, ${x_{s:n}}^{[B]}$.

A point prediction of the s-th order statistic $X_{s:n}$, $s = m+1, m+2, \ldots, n$, is $\sum_{j=1}^{B} x_{s:n}^{(j)}/B$ and a two-sided $100(1-\alpha)\%$ bootstrap prediction interval of $X_{s:n}$, $s = m+1, m+2, \ldots, n$, say $[X_{s:n}^{*L}, X_{s:n}^{*U}]$, is then given by

$$X_{s:n}^{*L} = {x_{s:n}}^{([B\alpha/2])}, \qquad X_{s:n}^{*U} = {x_{s:n}}^{([B(1-\alpha/2)])}.$$

We denote this method as *Boot*.

In Step 1, beside using the MLE, one can consider a bias-reduced estimator based on the MLE (Cox and Snell, 1968). The first-order bias of the MLE can be approximated as

$$Bias(\hat{\theta}) = k_2^{-2}(0.5k_3 + k_{21}),$$

where

$$\begin{aligned} k_{21} &= E\left[\left(\frac{\partial^2 \ln L(\theta)}{\partial\theta^2}\right)\left(\frac{\partial \ln L(\theta)}{\partial\theta}\right)\right], \\ k_2 &= E\left[\frac{\partial^2 \ln L(\theta)}{\partial\theta^2}\right], \\ k_3 &= E\left[\frac{\partial^3 \ln L(\theta)}{\partial\theta^3}\right], \end{aligned}$$

with

$$\begin{aligned} \frac{\partial \ln L(\theta)}{\partial\theta} &= \frac{2m}{\theta} - \frac{n}{(1+\theta)} + \frac{(n-m)(1+x_m)}{1+\theta+\theta x_m)} - \left[\sum_{i=1}^{m} x_i + (n-m)x_m\right], \\ \frac{\partial^2 \ln L(\theta)}{\partial\theta^2} &= -\frac{2m}{\theta^2} + \frac{n}{(1+\theta)^2} - \frac{(n-m)(1+x_m)^2}{1+\theta+\theta x_m)^2}, \\ \frac{\partial^3 \ln L(\theta)}{\partial\theta^3} &= \frac{4m}{\theta^3} - \frac{n}{(1+\theta)^3} + \frac{(n-m)(1+x_m)^3}{1+\theta+\theta x_m)^3}. \end{aligned}$$

Hence, the bias of $\hat{\theta}$ can be approximated by

$$\widehat{Bias}(\hat{\theta}) = \hat{k}_2^{-2}(0.5\hat{k}_3 + \hat{k}_{21}),$$

where

$$\hat{k}_{21} = \left(\frac{\partial^2 \ln L(\theta)}{\partial\theta^2}\right)\left(\frac{\partial \ln L(\theta)}{\partial\theta}\right)\Bigg|_{\theta=\hat{\theta}},$$

$$\hat{k}_2 = \frac{\partial^2 \ln L(\theta)}{\partial\theta^2}\Bigg|_{\theta=\hat{\theta}},$$

$$\hat{k}_3 = \frac{\partial^3 \ln L(\theta)}{\partial\theta^3}\Bigg|_{\theta=\hat{\theta}}.$$

A bias-corrected MLE can be obtained as $\tilde{\theta} = \hat{\theta} - \widehat{Bias}(\hat{\theta})$. Hence, the resampling method described above can be applied by using $\tilde{\theta}$ in place of $\hat{\theta}$ in Step 1. We denote this method as *BBR*.

7 Illustrative Example

To illustrate the prediction methods developed in this paper, we present an analysis of a real data set here. The following data set has been presented in Murthy *et al.* (2004). The data refer to the time between failures for some repairable items as

0.11	0.30	0.40	0.45	0.59	0.63	0.70	0.71	0.74	0.77
0.94	1.06	1.17	1.23	1.23	1.24	1.43	1.46	1.49	1.74
1.82	1.86	1.97	2.23	2.37	2.46	2.63	3.46	4.36	4.73

To check the validity of using Lindley distribution for fitting this data set, Kolmogorov-Smirnov (K-S) test is applied. The K-S statistic of the distance between the fitted and the empirical distribution functions (based on the parameter $\theta = 0.976$) is 0.561 and the corresponding p-value is 0.141. Therefore, it is reasonable to use the Lindley distribution to fit the data.

Suppose the life test is ended when the 25-th observation is observed, i.e., when we observe a Type-II censored sample with $n = 30$ and $m = 25$. Based on the prediction methods presented in this paper, we computed the point and interval predictions of the censored lifetime $Y = X_{s+25:30}$ ($s = 1, 2, 3, 4, 5$) and these results are presented in Table 1.

From Table 1, we see that various point predictors are quite close to the true observations. We can also observe that different prediction intervals considered here contain the true observations. Note that, as mentioned in Section 5, the distribution of Z given $\mathbf{X} = \mathbf{x}$ is a unimodal function of z, for $1 < s < 5$. Therefore, the HCD prediction method works here for all $Y = X_{s+25:30}$ ($s = 1, 2, 3, 4, 5$), except $s = 1$ and $s = 5$.

Table 1: Point predictors and 95% confidence prediction intervals for Y.

Point prediction						
	True Value	MLP	BUP	CMP	Boot	BBR
$s = 1$	2.46	2.370	2.636	2.557	2.636	2.643
$s = 2$	2.63	2.665	2.964	2.874	2.960	2.983
$s = 3$	3.46	3.037	3.392	3.283	3.383	3.424
$s = 4$	4.36	3.552	4.020	3.875	4.008	4.070
$s = 5$	4.73	4.414	5.237	4.969	5.221	5.327

Interval prediction					
	True Value	Pivot Method	HCD Method	Boot Method	BBR Method
$s = 1$	2.46	(2.376, 3.341)	—	(2.377, 3.474)	(2.377, 3.361)
$s = 2$	2.63	(2.443, 4.002)	(2.405, 3.812)	(2.442, 4.007)	(2.445, 4.024)
$s = 3$	3.46	(2.583, 4.817)	(2.584, 4.814)	(2.582, 4.782)	(2.587, 4.892)
$s = 4$	4.36	(2.815, 6.047)	(2.903, 6.883)	(2.805, 6.006)	(2.836, 6.115)
$s = 5$	4.73	(3.228, 8.781)	—	(3.222, 8.832)	(3.270, 8.871)

8 Monte Carlo Simulation Study

In this section, a Monte Carlo simulation study is employed to evaluate the performances of the different proposed predictors. We compare the performances of the point predictors MLP, BUP and CMP in terms of their biases and mean square prediction errors (MSPEs). The predictive intervals obtained by different methods are compared by means of their estimated average widths (AW) and coverage probabilities (CP).

For given (n, m, θ), we have randomly generated 10,000 sets of Type-II censored sample $X_{1:n}, ..., X_{m:n}$ from the Lindley distribution. Then, we have computed different point predictors for the sth future failure time $Y = X_{s+m:n} (s = 1, 2, \cdots, n-m)$. We also computed the 95% prediction intervals for Y. In our simulation, we have considered $n = 10, 15, 20$ and $\theta = 0.75, 1, 2$. Table 2 presents the biases and MSPEs of different point predictors. The biases and MSPEs are computed as follows. Suppose $\hat{Y}_i$ is the prediction of Y obtained in the i-th iteration of simulation, where $i = 1, \ldots, N = 10,000$, then we compute the bias and MSPE as

$$Bias = \frac{1}{N}\sum_{i=1}^{N}(\hat{Y}_i - Y)$$

and

$$MSPE = \frac{1}{N}\sum_{i=1}^{N}(\hat{Y}_i - Y)^2.$$

The average confidence lengths and the corresponding coverage probabilities of 95% prediction intervals are also computed and reported in Table 3. If (L_i, U_i) is the 95% prediction interval of Y obtained in the i-th iteration of simulation, where $i = 1, \ldots, N = 10,000$, then the coverage probability (CP) is computed as

$$\text{CP} = \frac{1}{10000}\sum_{i=1}^{10000} I\{L_i < Y < U_i\},$$

where $I\{\cdot\}$ denotes the indicator function.

From Table 2, it is observed that, in most cases, the BUP and CMP perform better than other point predictors in terms of MSPEs. Since the BUP is nearly unbiased, one may prefer to use the BUP compared to the CMP. The Boot predictor works better than the BBR predictor in general. We also observe that the MLP does not work well because it gives the largest biases and MSPEs among all the predictors considered here. This point can be also observed in the papers by Awad and Raqab (2000), Basak *et al.* (2006) and Asgharzadeh *et al.* (2015). For fixed values of n and m, the biases and MSREs increase as s increases.

From Table 3, we can see that the Pivot method and the BBR prediction intervals have simulated coverage probabilities closer to the 95% nominal level in most cases. However, the trade-off is that the Pivot method and BBR prediction intervals have larger average widths. The simulated average widths of prediction intervals based on the Pivot method are smaller than that of the BBR prediction intervals. For fixed values of n and m, the average widths of different prediction intervals increase as s increases.

Overall speaking, based on the simulation results here, we would recommend using the best unbiased predictor for point prediction of the future failures and using the prediction interval based on the Pivot method for interval prediction. Note that the methodology presented in this work could be extended to Type-I censored data with some straightforward modifications. Specifically, suppose that a Type-I censored life testing experiment with n units gets terminated at time τ and m failures are observed before time τ while $(n - m)$ items are censored. Here, τ is a pre-fixed value and m is a random variable. Conditional on the observed value of m, the prediction methods presented in this paper can be used by setting $x_m = \tau$.

Table 2: Biases and MSPEs of point predictors.

	n	m	s		MLP	BUP	CMP	Boot	BBR
$\theta = 0.75$	10	7	1	Bias	-0.595	-0.003	-0.175	-0.029	0.031
				MSPE	0.679	0.364	0.377	0.386	0.392
			2	Bias	-0.780	-0.020	-0.237	-0.028	0.113
				MSPE	1.647	1.192	1.196	1.208	1.250
			3	Bias	-1.308	-0.045	-0.414	-0.024	0.263
				MSPE	2.450	2.273	2.281	4.271	4.463
	15	10	1	Bias	-0.359	-0.005	-0.103	-0.002	0.025
				MSPE	0.253	0.134	0.140	0.136	0.143
			2	Bias	-0.420	-0.007	-0.127	-0.003	0.029
				MSPE	0.506	0.363	0.367	0.365	0.383
			3	Bias	-0.521	-0.024	-0.164	-0.007	0.056
				MSPE	0.958	0.760	0.763	0.758	0.815
			4	Bias	-0.697	-0.037	-0.221	-0.023	0.088
				MSPE	1.914	1.582	1.586	1.592	1.681
			5	Bias	-1.160	-0.011	-0.353	-0.007	0.184
				MSPE	2.749	2.419	2.424	4.452	4.666
	20	15	1	Bias	-0.352	-0.001	-0.105	-0.006	0.010
				MSPE	0.240	0.122	0.130	0.131	0.132
			2	Bias	-0.402	-0.004	-0.123	-0.013	0.022
				MSPE	0.473	0.332	0.340	0.334	0.337
			3	Bias	-0.484	-0.009	-0.149	-0.012	0.047
				MSPE	0.877	0.689	0.696	0.681	0.689
			4	Bias	-0.644	-0.018	-0.203	-0.008	0.085
				MSPE	1.780	1.465	1.473	1.423	1.442
			5	Bias	-1.083	-0.011	-0.332	-0.032	0.124
				MSPE	2.031	1.919	1.928	3.917	3.964
	n	m	s		MLP	BUP	CMP	Boot	BBR
$\theta = 1.0$	10	7	1	Bias	-0.442	-0.003	-0.132	-0.005	0.042
				MSPE	0.376	0.200	0.209	0.187	0.192
			2	Bias	-0.584	-0.016	-0.178	-0.005	0.108
				MSPE	0.911	0.653	0.656	0.646	0.675
			3	Bias	-1.006	-0.055	-0.332	-0.015	0.216
				MSPE	1.643	1.444	1.449	2.431	2.551
	15	10	1	Bias	-0.270	-0.002	-0.082	0.000	0.020
				MSPE	0.144	0.077	0.081	0.074	0.075
			2	Bias	-0.318	-0.009	-0.099	-0.016	0.027
				MSPE	0.283	0.201	0.205	0.207	0.210
			3	Bias	-0.392	-0.018	-0.123	-0.026	0.047
				MSPE	0.550	0.439	0.044	0.443	0.452
			4	Bias	-0.514	-0.015	-0.154	-0.047	0.066
				MSPE	1.080	0.903	0.906	0.942	0.961
			5	Bias	-0.882	-0.020	-0.276	-0.011	0.178
				MSPE	2.075	1.537	1.542	2.416	2.494
	20	15	1	Bias	-0.265	-0.005	-0.083	-0.002	0.010
				MSPE	0.137	0.071	0.076	0.071	0.132
			2	Bias	-0.300	-0.004	-0.093	-0.002	0.022
				MSPE	0.257	0.178	0.182	0.176	0.337
			3	Bias	-0.376	-0.021	-0.126	-0.010	0.047
				MSPE	0.501	0.388	0.394	0.386	0.689
			4	Bias	-0.483	-0.012	-0.151	0.002	0.085
				MSPE	0.978	0.803	0.807	0.803	1.442
			5	Bias	-0.839	-0.016	-0.272	0.010	0.124
				MSPE	1.828	1.276	1.277	2.202	3.964
	n	m	s		MLP	BUP	CMP	Boot	BBR
$\theta = 2.0$	10	7	1	Bias	-0.208	-0.000	-0.062	-0.006	0.022
				MSPE	0.085	0.048	0.049	0.048	0.050
			2	Bias	-0.294	-0.014	-0.095	-0.003	0.065
				MSPE	0.230	0.165	0.166	0.157	0.166
			3	Bias	-0.488	-0.011	-0.149	-0.013	0.129
				MSPE	0.746	0.591	0.595	0.610	0.650
	15	10	1	Bias	-0.127	-0.000	-0.038	-0.002	0.010
				MSPE	0.033	0.018	0.029	0.017	0.017
			2	Bias	-0.152	-0.002	-0.046	-0.008	0.018
				MSPE	0.063	0.045	0.055	0.048	0.049
			3	Bias	-0.190	-0.005	-0.057	-0.021	0.022
				MSPE	0.126	0.100	0.107	0.107	0.110
			4	Bias	-0.250	-0.000	-0.069	-0.031	0.037
				MSPE	0.248	0.209	0.217	0.221	0.227
			5	Bias	-0.456	-0.019	-0.147	-0.028	0.088
				MSPE	0.772	0.634	0.638	0.593	0.615
	20	15	1	Bias	-0.125	-0.001	-0.038	0.003	0.011
				MSPE	0.031	0.016	0.017	0.017	0.017
			2	Bias	-0.147	-0.003	-0.047	0.002	0.019
				MSPE	0.061	0.043	0.044	0.041	0.042
			3	Bias	-0.175	-0.000	-0.052	0.005	0.033
				MSPE	0.111	0.087	0.087	0.086	0.088
			4	Bias	-0.244	-0.009	-0.078	0.000	0.045
				MSPE	0.236	0.191	0.192	0.193	0.198
			5	Bias	-0.409	-0.008	-0.121	0.005	0.081
				MSPE	0.695	0.571	0.573	0.547	0.560

Table 3: Simulated average widths (AW) and coverage probabilities (CP) of different 95% predictive intervals.

	n	m	s		Pivot Method	HCD Method	Boot Method	BBR Method
$\theta = 0.75$	10	7	1	AW	2.108	—	2.065	2.265
				CP	0.930	—	0.919	0.930
			2	AW	3.751	3.717	3.692	4.032
				CP	0.925	0.929	0.909	0.921
			3	AW	6.979	—	6.913	7.520
				CP	0.916	—	0.913	0.926
	15	10	1	AW	1.305	—	1.275	1.348
				CP	0.938	—	0.933	0.936
			2	AW	1.830	1.814	2.052	2.170
				CP	0.925	0.921	0.919	0.924
			3	AW	2.956	2.951	2.924	3.092
				CP	0.916	0.915	0.913	0.916
			4	AW	4.247	4.208	4.195	4.419
				CP	0.912	0.911	0.908	0.913
			5	AW	6.239	—	7.138	7.514
				CP	0.909	—	0.913	0.913
	20	15	1	AW	1.267	—	1.228	1.283
				CP	0.944	—	0.936	0.939
			2	AW	2.043	1.845	1.993	2.079
				CP	0.931	0.933	0.929	0.936
			3	AW	2.923	2.918	2.854	2.975
				CP	0.925	0.927	0.918	0.921
			4	AW	3.197	3.158	4.116	4.284
				CP	0.924	0.926	0.924	0.931
			5	AW	5.169	—	7.040	7.318
				CP	0.923	—	0.925	0.928
$\theta = 1.0$	10	7	1	AW	1.566	—	1.533	1.693
				CP	0.926	—	0.922	0.936
			2	AW	2.785	2.705	2.748	3.023
				CP	0.923	0.926	0.909	0.922
			3	AW	5.211	—	5.160	5.652
				CP	0.922	—	0.910	0.922
	15	10	1	AW	0.963	—	1.275	1.348
				CP	0.931	—	0.933	0.936
			2	AW	1.550	1.401	2.052	2.170
				CP	0.919	0.917	0.919	0.924
			3	AW	2.207	2.203	2.924	3.092
				CP	0.911	0.910	0.913	0.916
			4	AW	3.172	3.094	4.195	4.419
				CP	0.909	0.908	0.908	0.913
			5	AW	4.381	—	7.138	7.514
				CP	0.906	—	0.913	0.913
	20	15	1	AW	0.914	—	0.912	1.283
				CP	0.940	—	0.942	0.939
			2	AW	1.518	1.370	1.482	2.079
				CP	0.928	0.931	0.930	0.936
			3	AW	2.164	2.160	2.126	2.975
				CP	0.924	0.926	0.911	0.921
			4	AW	2.937	2.857	3.071	4.284
				CP	0.923	0.935	0.926	0.931
			5	AW	3.355	—	5.261	7.318
				CP	0.921	—	0.918	0.928
$\theta = 2.0$	10	7	1	AW	0.748	—	0.737	0.834
				CP	0.925	—	0.920	0.933
			2	AW	1.349	1.286	1.334	1.504
				CP	0.920	0.924	0.906	0.917
			3	AW	2.547	—	2.535	2.845
				CP	0.927	—	0.896	0.904
	15	10	1	AW	0.459	—	0.440	0.480
				CP	0.930	—	0.929	0.933
			2	AW	0.745	0.672	0.718	0.782
				CP	0.918	0.915	0.903	0.921
			3	AW	1.066	1.064	1.035	1.125
				CP	0.910	0.910	0.894	0.906
			4	AW	1.542	1.900	1.497	1.624
				CP	0.908	0.907	0.900	0.909
			5	AW	2.644	—	2.582	2.793
				CP	0.904	—	0.913	0.919
	20	15	1	AW	0.450	—	0.438	0.464
				CP	0.939	—	0.929	0.933
			2	AW	0.732	0.660	0.716	0.757
				CP	0.925	0.929	0.926	0.931
			3	AW	1.052	1.050	1.033	1.092
				CP	0.922	0.921	0.928	0.929
			4	AW	1.529	1.486	1.501	1.585
				CP	0.920	0.917	0.903	0.913
			5	AW	2.645	—	2.592	2.731
				CP	0.918	—	0.920	0.925

Acknowledgements

The authors would like to thank the referees for their valuable suggestions and comments which improved the paper. Research of the second author was supported by a grant from the University of Mazandaran.

References

Ali, S., Aslam, M., and Kazmi, S. M. A. (2013). A study of the effect of the loss function on Bayes estimate, posterior risk and hazard function for Lindley distribution. *Applied Mathematical Modelling*, **37**, 6068-6078.

Al-Mutairi D, K., Ghitany, M. E., and Kundu, D. (2013). Inference on stress strength reliability from Lindley distribution. *Communication in Statistics–Theory and Methods*, **42**, 1443-1463.

Asgharzadeh, A., Bakouch, H. S., Nadarajah, S., and Sharafi, F. (2016a). A new weighted Lindley distribution with application. *Brazilian Journal of Probability and Statistics*, **30**(1), 1-27. DOI: 10.1214/14-BJPS253.

Asgharzadeh, A., Fallah, A., Raqab, M. Z,. and Valiollahi, R. (2016b). Statistical inference based on Lindley record data. *Statistical Papers*, DOI:10.1007/s00362-016-0788-1.

Asgharzadeh., A, Valiollahi, R., and Kundu, D. (2015). Prediction for future failures in Weibull distribution under hybrid censoring. *Journal of Statistical Computation and Simulation*, **85** (4), 824-838.

Awad, A M., and Raqab, M. Z. (2000). Prediction intervals for the future record values from exponential distribution: comparative study. *Journal of Statistical Computation and Simulation*, **65**, 325-340.

Bakouch, H. S., Al-Zahrani, B. M., Al-Shomrani, A. A., Marchi, V. A. A., and Louzada, F. (2012). An extended Lindley distribution. *Journal of the Korean Statistical Society*, **41**, 75-85.

Basak, I., Basak, P., and Balakrishnan, N. (2006). On some predictors of times to failure of censored items in progressively censored samples. *Computational Statistics and Data Analysis*, **50**, 1313-1337.

Chapeau-Blondeau, F., and Monir, A. (2002). Evaluation of the Lambert W function and application to generation of generalized Gaussian noise with exponent 1/2. *IEEE Transactions on Signal Processing*, **50**(9), 2160-2165.

Efron, B., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*, Chapman & Hall, Boca Raton, Florida.

Ghitany, M. E., Al-Mutairi, D. K., Balakrishnan, N., and Al-Enezi, L. J. (2013). Power Lindley distribution and associated inference. *Computational Statistics and Data Analysis*, **64**, 20-33.

Ghitany, M. E., Alqallaf, F., Al-Mutairi, D. K., and Husain, H. A. (2011). A two-parameter weighted Lindley distribution and its applications to survival data. *Mathematics and Computers in Simulation*, **81**, 1190-1201.

Ghitany, M. E., Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, **78**, 493-506.

Gomez-Deniz, E., Sordo, M. A., and Calderin-Ojeda, E. (2014). The log-Lindley distribution as an alternative to the beta regression model with applications in insurance. *Insurance: Mathematics and Economics*, **54**, 49-57.

Gupta, P.,K., and Singh, B. (2013). Parameter estimation of Lindley distribution with hybrid censored data. *International Journal of System Assurance Engineering and Management*, **4**(4), 378-385.

Kaminsky, K. S., and Rhodin, L. S. (1985). Maximum likelihood prediction. *Annals of the Institute of Statistical Mathematics*, **37**, 507-517.

Krishna, H., and Kumar, K. (2011). Reliability estimation in Lindley distribution with progressively Type II right censored sample. *Mathematics and Computers in Simulation*, **82**(2), 281-294.

Lindley, D. V. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society, A*, **20**, 102-107.

Murthy, D. N. P., Xie, M., and Jiang, R. (2004). *Weibull Models*, Wiley series in probability and statistics. NJ: John Wiley and Sons.

Nadarajah, S., Bakouch, H. S., and Tahmasbi, R. (2011). A generalized Lindley distribution. *Sankhyā, B*, **73**, 331-359.

Raqab, M. Z., and Nagaraja, H. N. (1995). On some predictors of future order statistics, *Metron*, **53**(12), 185-204.

Zakerzadeh, H., and Dolati, A. (2009). Generalized Lindley distribution. *Journal of Mathematical Extension*, **3**, 13-25.