Prediction of Future Lifetimes for a Simple Step-Stress Model with Type-II Censoring and Rayleigh Distribution

MOHAMMAD A. AMLEH
Department of Mathematics, Faculty of Science
Zarqa University, Zarqa, 13110
JORDAN

Abstract: - In this paper, we discuss the prediction problem of the lifetimes to failure of units from Rayleigh distribution with Type-II censoring for a simple step-stress setup under cumulative exposure model. We consider several methods of point prediction, including maximum likelihood predictor, conditional median predictor, and best unbiased predictor. In addition, we discuss the prediction intervals for future lifetimes of the censored units using pivotal quantity, highest conditional density, and shortest-length based methods. Monte Carlo simulation is conducted to compare the proposed prediction methods. Further, a real data set is analyzed for illustrative purposes.

Keywords: - Conditional median predictor; Cumulative exposure model; Maximum likelihood prediction; Pivot quantity; Prediction intervals; Rayleigh distribution; Step-stress accelerated life test.

Received: June 8, 2021. Revised: February 3, 2022. Accepted: February 20, 2022. Published: March 24, 2022.

1 Introduction

Accelerated life tests (ALTs) are commonly used to test components operated at higher than usual levels of stress. The failure data obtained from such tests are then transformed to estimate the distribution of failures under specified conditions, which improves component designs and makes better component selections.

In ALT, the model is chosen according to the relationship between the parameters of the lifetime distribution and the conditions of the accelerated stress. If we use a constant stress level and some selected stress levels are very low, then we get many non-failed units during the testing time, which reduces the effectiveness of the test. To overcome this problem, step stress accelerated life test (SSALT) can be used. For more details on ALTs, one may refer to Nelson [20] and Kundu and Ganguly [17].

In the SSALT, the level of stress in the test will be increased in steps at different intermediate stages of the experiment. Accordingly, a test unit is subjected to a specified level of stress for a prefixed period of time, if it does not fail during that period of time, then the stress level is changed for future prefixed time. This process continues until the test units fail or some termination conditions will be used. If there are two levels of stress, the SSALT is known as simple SSALT. In order to analyze the data under SSALT, there is more than one model that connects the lifetime's distribution under various stress levels to the failure times under the step stress test. The most popular model is known as the cumulative exposure model (CEM), which was proposed by Nelson [19]. In this model, it is assumed that the remaining lifetime of the experiment units is dependent only on the cumulative exposure the units have experienced, with no memory on how the exposure was accumulated. Statistical inferences of step-stress test under CEM were discussed by many authors. Estimation of the parameters in a simple step-stress test under CEM for Weibull and exponential distributions are addressed by Bai and Kim [4] and Xiong [22], respectively. Balakrishnan et al. [5] presented a simple step-stress model under Type-I censoring and lognormally distributed lifetimes. Mitra et al. [18] discussed a simple step-stress model for two-parameter exponential distribution with Type-II censoring.

For Rayleigh distribution, Ebrahem and Al-Masri [11] discussed the estimation problem of the parameters for a simple step-stress model of Rayleigh distribution with log-linear link function. Chandra and Khan [10] presented the estimation problem of the parameters for simple step-stress model under Rayleigh distribution with Type-I and Type-II censoring. Kumar et al [16] considered the Bayesian inference for Rayleigh distribution under step-stress partially accelerated test with progressive Type-II censoring and binomial removal. Kotb and El-Din [15] presented a parametric inference for step-stress tests from Rayleigh distribution under ordered ranked set sampling.

It may worth mentioning that no attention has been paid to the problem of prediction of new lifetimes of Rayleigh distribution under CEM. In fact, the prediction problem has not been discussed extensively for step-stress model in the literature. Basak [6], and Basak and Balakrishnan ([7], [8]) considered the problem of predicting failure times of censored units for a simple step-stress model from exponential distribution with Type-I censoring and Type-II censoring, respectively. Recently, Amleh and Raqab ([2], [3]) discussed the prediction problem for step-stress plan for Lomax distribution under CEM, and for Weibull distribution under Khamis-Higgens model, respectively.

In this paper, the simple SSALT for the Rayleigh distribution based on CEM is considered. It is assumed that failures occur according to Type-II censoring scheme, in which the experiment is terminated as soon as the r^{th} failure occurs. Specifically, the aim of the paper is predicting future order statistics based on Type-II censored units under simple step-stress setup with Rayleigh CEM using point prediction and prediction intervals.

The rest of the paper is organized as follows. The CEM under Rayleigh distribution, basic model assumptions and maximum likelihood estimation of the original parameters based on the observed data are discussed in Section 2. Point predictors including maximum likelihood predictor, conditional median predictor, and the best unbiased predictor are presented in Section 3. In Section 4, we develop different methods for obtaining prediction limits of the censored lifetimes. To assess the effectiveness of the prediction procedures, we perform a simulation study and real data analysis in Section 5. Finally, we conclude the paper in Section 6.

2 Model description and maximum likelihood estimation

Rayleigh distribution was introduced in 1880 as part of a problem in the field of acoustics. Over the following years, significant work has taken place in the distribution in different fields of science and technology. Rayleigh distribution is related to other known distributions such as Weibull, chi-square and extreme values distributions. An important feature of the Rayleigh distribution is that its hazard rate function is an increasing function of time. This means that if the failure times have Rayleigh distribution, an intense aging item occurs. For more details on Rayleigh distribution one may refer to Johnson et al. [13]. The probability density function (pdf) of the Rayleigh distribution is given by

$$f(t,\theta)=\frac{t}{\theta^2}e^{-\frac{t^2}{2\theta^2}}, t>0\,,\theta>0, \qquad (1)$$

with cumulative distribution function (cdf)

$$F(t,\theta)=1-e^{-\frac{t^2}{2\theta^2}}, t>0\,,\theta>0, \qquad (2)$$

where θ is the scale parameter. The hazard rate function of the Rayleigh distribution is increasing in t and given by

$$h(t)=\frac{t}{\theta^2},$$

So, Rayleigh distribution may describe the lifetime of an increasing failure rate items.

The simple step-stress test under Type-II censoring is performed as follows. All n units are initially put on the lower stress S_1 and run until time τ. Then, the stress is increased to high level S_2, and the test continues until a pre-determined number of failures r are observed. Let n_1 denotes the random number of failures before τ, and $n_2=r-n_1$, denotes the number of failures after τ. If $n_1=r$, then the test is terminated at the first level. Otherwise, the stress level is accelerated to the next step, and the test continues until the required r failures. The following are the basic assumptions that specify our model:

1- Units are tested at two levels of stress $S_1<S_2$;
2- The lifetimes of the units for both stress levels follow Rayleigh distribution;
3- The scale parameters for the lifetime distribution are $\theta_j, j=1,2$, corresponding to stress level $S_j, j=1,2$;
4- Failures follow the CEM.

According to the above assumptions, the ordered lifetimes that are observed, which are denoted by the vector data $\boldsymbol{t}$, have the following form

$$t_{1:n}<\cdots<t_{n_1:n}<\tau\le t_{n_1+1:n}<\cdots<t_{r:n}. \qquad (3)$$

Here, $\boldsymbol{t}$ represents the observed values of the variable $\boldsymbol{T}=(T_{1:n},\dots,T_{n_1},T_{n_1+1:n},\cdots,T_r)$, which denotes the Type-II censored order statistics. The CEM for simple step-stress test is given by

$$F(t)=\begin{cases}F_1(t), & 0\le t<\tau\\ F_2(t-\tau+h)\,, & \tau\le t<\infty,\end{cases} \qquad (4)$$

where the equivalent shifting time, h, is a solution of the equation $F_1(\tau)=F_2(h)$. By solving the above equation for h, we have $h = \frac{\theta_2}{\theta_1}\tau$. As a result of that, the Rayleigh CEM for simple step-stress test is distributed as

$$F(t) = \begin{cases} 1 - e^{-\frac{t^2}{2\theta_1^2}}, & 0 \le t < \tau \\ 1 - e^{-\frac{\left(\frac{t-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}{2}}, & \tau \le t < \infty, \end{cases} \quad (5)$$

with the corresponding pdf

$$f(t) = \begin{cases} \frac{t}{\theta_1^2}\, e^{-\frac{t^2}{2\theta_1^2}}, & 0 \le t < \tau \\ \frac{1}{\theta_2}\left(\frac{t-\tau}{\theta_2} + \frac{\tau}{\theta_1}\right) e^{-\frac{\left(\frac{t-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}{2}}, & \tau \le t < \infty. \end{cases} \quad (6)$$

The likelihood function of the parameters θ_1 and θ_2 based on the observed Type-II censored data $\boldsymbol{t}$ is given by

$$L(\theta_1, \theta_2 | \boldsymbol{t}) = \begin{cases} \frac{n!}{r!}\prod_{i=1}^{n_1} f_1(t_{i:n})\,[1 - F_1(t_{r:n})]^{n-r}, \ n_1 = r & (7\,a) \\ \frac{n!}{r!}\prod_{i=1}^{r} f_2(t_{i:n})\,[1 - F_2(t_{r:n})]^{n-r}, \ n_1 = 0 & (7\,b) \\ \frac{n!}{r!}\prod_{i=1}^{r} f(t_{i:n})\,[1 - F(t_{r:n})]^{n-r}, \ 1 \le n_1 \le r-1, & (7\,c) \end{cases}$$

Based on the likelihood function given in (7 a), (7 b) and (7 c), it is observed that the maximum likelihood estimators (MLEs) of the parameters θ_1 and θ_2 exist only if $1 \le n_1 \le r - 1$. Therefore, according to the step-stress setup, the likelihood function in (2.7 c) is given by

$$L(\theta_1, \theta_2 | \boldsymbol{t}) = \frac{n!}{n_1!\,(r-n_1)!}\prod_{i=1}^{n_1} f_1(t_{i:n}) \prod_{i=n_1+1}^{r} f_2(t_{i:n}) \times [1 - F_2(t_{r:n})]^{n-r}. \quad (8)$$

Using Eq.'s (5) and (6), we have

$$L(\theta_1, \theta_2 | \boldsymbol{t}) \propto \prod_{i=1}^{n_1}\left\{\frac{t_{i:n}}{\theta_1^2}\, e^{-\frac{t_{i:n}^2}{2\theta_1^2}}\right\} \times \prod_{i=n_1+1}^{r}\left\{\frac{1}{\theta_2}\left(\frac{t_{i:n}-\tau}{\theta_2} + \frac{\tau}{\theta_1}\right) e^{-\frac{\left(\frac{t_{i:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}{2}}\right\} \times \left[e^{-\frac{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}{2}}\right]^{n-r}, \quad (9)$$

which can be simplified as:

$$L(\theta_1, \theta_2 | \boldsymbol{t}) \propto \theta_1^{-2n_1}\theta_2^{-n_2}\prod_{i=1}^{n_1} t_{i:n}\prod_{i=n_1+1}^{r}\left(\frac{t_{i:n}-\tau}{\theta_2} + \frac{\tau}{\theta_1}\right) \times e^{-\frac{1}{2}\left[\frac{1}{\theta_1^2}\sum_{i=1}^{n_1} t_{i:n}^2 + \sum_{i=n_1+1}^{r}\left(\frac{t_{i:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2 + (n-r)\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]}. \quad (10)$$

Consequently, the log-likelihood function $L^* = \log(L)$ is given by

$$L^*(\theta_1, \theta_2 | \boldsymbol{t}) \propto -2n_1\log\theta_1 - n_2\log\theta_2 + \sum_{i=1}^{n_1}\log\, t_{i:n} + \sum_{i=n_1+1}^{r}\log\left(\frac{t_{i:n}-\tau}{\theta_2} + \frac{\tau}{\theta_1}\right) - \frac{1}{2}\left[\frac{1}{\theta_1^2}\sum_{i=1}^{n_1} t_{i:n}^2 + \sum_{i=n_1+1}^{r}\left(\frac{t_{i:n}-\tau}{\theta_2} + \frac{\tau}{\theta_1}\right)^2 + (n-r)\left(\frac{t_{r:n}-\tau}{\theta_2} + \frac{\tau}{\theta_1}\right)^2\right]. \quad (11)$$

So, the likelihood equations are given by

$$\frac{\partial L^*}{\partial\theta_1} = \frac{1}{\theta_1^3}\sum_{i=1}^{n_1} t_{i:n}^2 - \frac{\theta_2}{\theta_1}\sum_{i=n_1+1}^{r}\frac{\tau}{\theta_1(t_{i:n}-\tau) + \theta_2\tau} + \frac{\tau}{\theta_1^2}\left[\sum_{i=n_1+1}^{r}\left(\frac{t_{i:n}-\tau}{\theta_2} + \frac{\tau}{\theta_1}\right) + (n-r)\left(\frac{t_{r:n}-\tau}{\theta_2} + \frac{\tau}{\theta_1}\right)\right] - \frac{2n_1}{\theta_1} = 0. \quad (12)$$

$$\frac{\partial L^*}{\partial\theta_2} = -\frac{n_2}{\theta_2} + (n-r)\left(\frac{t_{r:n}-\tau}{\theta_2} + \frac{\tau}{\theta_1}\right)\frac{(t_{r:n}-\tau)}{\theta_2^2} + \sum_{i=n_1+1}^{r}\left\{-\frac{\theta_1}{\theta_2}.\frac{(t_{i:n}-\tau)}{\theta_1(t_{i:n}-\tau)+\theta_2\tau} + \left(\frac{t_{i:n}-\tau}{\theta_2} + \frac{\tau}{\theta_1}\right)\frac{(t_{i:n}-\tau)}{\theta_2^2}\right\} = 0. \quad (13)$$

The estimation procedure, through equations (12) and (13), does not result in closed form. Therefore, Eq.'s (12) and (13) can be solved simultaneously using a numerical technique as Newton-Raphson method, or similar methods, see [1]. The algorithm used for generating the data and computing the MLEs of the parameters θ_1 and θ_2 is performed according to the following algorithm:

Step 1: Generate a random sample of size n following standard uniform distribution $\mathrm{U}(0,1)$, and obtain the order statistics:

$$U_{1:n} < U_{2:n} < \cdots < U_{n:n};$$

Step 2: Find the random variable n_1 such that $U_{n_1} < P(T \le \tau) = F_1(\tau) \le U_{n_1+1:n}$, where T represents the failure time, so we have:

$$U_{n_1} < 1 - e^{-\frac{\tau^2}{2\theta_1^2}} \le U_{n_1+1:n}.$$

Step 3: Generate the data based on the order statistics $U_{i:n}$ as follows:

$$t_{i:n} = \begin{cases} \theta_1\sqrt{-2\log(1-U_{i:n})}\,, & i = 1,2,\dots,n_1 \\ \theta_2\sqrt{-2\log(1-U_{i:n})} + \tau\left(1-\frac{\theta_2}{\theta_1}\right), & i = n_1+1,\dots,r \end{cases} \quad (14)$$

Step 4: Compute the MLEs of θ_1 and θ_2 using Eq.'s (12) and (13) based on the censored data

$$t_{i:n}, t_{2:n}, \dots, t_{n_1:n}, t_{n_1+1:n}, \dots, t_{r:n},$$

as in (14).

3 Prediction of future order statistics

In this section, we discuss the problem of predicting new failure times based on some observed Rayleigh failure times under the CEM. The problem can be described as follows. Let $T_{1:n} < T_{2:n} < \cdots < T_{r:n}$ denote the observed ordered lifetime units, which is known as informative sample, and let $T_{s:n}$, $s = r+1,\dots,n$, be the unobserved future lifetime taken from the same sample. The prediction problem concerns on how we can predict the future lifetimes $T_{s:n}$, given the observed ordered statistics $T_{i:n}$, $0 < i \le r$.

Based on the Markovian property of censored order statistics, it is known that the conditional distribution of $Y = T_{s:n}$ given $\boldsymbol{T} = \boldsymbol{t}$, where:

$$\boldsymbol{t} = \left(t_{1:n}, \dots, t_{n_1:n}, t_{n_1+1:n}, \dots, t_{r:n}\right),$$

is equivalent to the distribution of $Y = T_{s:n}$ given $T_{r:n} = t_{r:n}$. Therefore, the density of Y given $\boldsymbol{T} = \boldsymbol{t}$ is the same as the density of the $(s-r)^{th}$ order statistic out of $(n-r)$ units from the population with left truncated density $\varphi(y) = \frac{f(y)}{1-F(t_{r:n})}, y > t_{r:n}$, where $F(y)$ and $f(y)$ are given in Section 2 as in Eq. (5) and (6), respectively. Therefore, the density of $Y = T_{s:n}$ given $\boldsymbol{T} = \boldsymbol{t}$ can be expressed as:

$$g_{T_{s:n}|\boldsymbol{T}}(y|\theta_1,\theta_2,data) = \frac{c}{\theta_2}\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right) \times \left\{1 - e^{-\frac{1}{2}\left[\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2 - \left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]}\right\}^{s-r-1} \times e^{-\frac{1}{2}(n-s+1)\left[\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2 - \left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]}, y > t_{r:n}, \quad (15)$$

where $c = \frac{(n-r)!}{(s-r-1)!(n-s)!}$.

3.1 Maximum likelihood predictor

The maximum likelihood prediction method was suggested by Kaminsky and Rhodin [12]. This method includes the prediction of future order statistics in addition to estimating the unknown parameters in the proposed model. The predictive likelihood function (PLF) of $Y = T_{s:n}$ is given by

$$l(y,\theta_1,\theta_2|\boldsymbol{t}) = l = g_{T_{s:n}|\boldsymbol{T}}(y|\boldsymbol{t},\theta_1,\theta_2).\,g_{\boldsymbol{T}}(\boldsymbol{t},\theta_1,\theta_2) = g_{T_{s:n}|T_{r:n}}(y|t_{r:n},\theta_1,\theta_2).\,g_{\boldsymbol{T}}(\boldsymbol{t},\theta_1,\theta_2), \quad (16)$$

where $g_{T_{s:n}|T_{r:n}}(y|t_{r:n},\theta_1,\theta_2)$ is the conditional density of $T_{s:n}$ given the observed value of $\boldsymbol{T} = \boldsymbol{t}$, as in Eq. (15), and $g_{\boldsymbol{T}}(\boldsymbol{t},\theta_1,\theta_2)$ is the density of $\boldsymbol{T}$**.** In fact, the PLF of $Y = T_{s:n}$ can be formed as

$$l \propto \prod_{i=1}^{n_1} f_1(t_{i:n}) \prod_{i=n_1+1}^{r} f_2(t_{i:n}) \times [F_2(y) - F_2(t_{r:n})]^{s-r-1} f_2(y)[1-F_2(y)]^{n-s}$$

$$0 \le n_1 \le r, r+1 \le s \le n. \quad (17)$$

Taking the case when $1 \le n_1 < r \le n$, we obtain

$$l \propto \theta_1^{-2n_1}\theta_2^{-(n_2+1)}\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)\prod_{i=1}^{n_1} t_{i:n} \times e^{-\frac{1}{2}\left[\frac{1}{\theta_1^2}\sum_{i=1}^{n_1} t_{i:n}^2 + \sum_{i=n_1+1}^{r}\left(\frac{t_{i:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2 + (n-s+1)\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]} \times \left[e^{-\frac{1}{2}\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2} - e^{-\frac{1}{2}\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}\right]^{s-r-1} \times \prod_{i=n_1+1}^{r}\left(\frac{t_{i:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right). \quad (18)$$

So, the log PLF can be written as

$$\log l \propto -2n_1\log\theta_1 - (n_2+1)\log\theta_2 + \sum_{i=1}^{n_1}\log\, t_{i:n} + \sum_{i=n_1+1}^{r}\log\left(\frac{t_{i:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right) + (s-r-1)\log\left[e^{-\frac{1}{2}\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2} - e^{-\frac{1}{2}\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}\right] - \frac{1}{2}\left[\frac{1}{\theta_1^2}\sum_{i=1}^{n_1} t_{i:n}^2 + \sum_{i=n_1+1}^{r}\left(\frac{t_{i:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2 + (n-s+1)\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right] + \log\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right). \quad (19)$$

Using (19), the predictive likelihood equations (PLEs) for θ_1, θ_2 and y are obtained and presented as follows:

$$\frac{\partial\, log\, l}{\partial\theta_1} = \frac{1}{\theta_1^3}\sum_{i=1}^{n_1}(t_{i:n}^2) - \frac{2n_1}{\theta_1} + \frac{\tau}{\theta_1^2}(s-r-1) \times$$

$$\left\{\frac{\left[\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right]e^{-\frac{1}{2}\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}-\left[\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right]e^{-\frac{1}{2}\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}}{e^{-\frac{1}{2}\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}-e^{-\frac{1}{2}\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}}\right\}$$

$$-\frac{\theta_2}{\theta_1}\left[\left(\frac{\tau}{\tau\theta_2+(y-\tau)\theta_1}\right)-\sum_{i=n_1+1}^{r}\left(\frac{\tau}{\tau\theta_2+(t_{i:n}-\tau)\theta_1}\right)\right]$$

$$\frac{\tau}{\theta_1^2}\left[\sum_{i=n_1+1}^{r}\left(\frac{t_{i:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)+(n-s+1)\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)\right]=0. \tag{20}$$

$$\frac{\partial\log l}{\partial\theta_2}=-\frac{n_2+1}{\theta_2}+(n-s+1)\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)\left(\frac{y-\tau}{\theta_2^2}\right)$$
$$-\frac{\theta_1}{\theta_2}\left(\frac{y-\tau}{\tau\theta_2+(y-\tau)\theta_1}\right)+\frac{(s-r-1)\theta_2^{-2}}{e^{-\frac{1}{2}\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}-e^{-\frac{1}{2}\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}}\times$$
$$\left\{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)(t_{r:n}-\tau)e^{-\frac{1}{2}\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}-\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)(y-\tau)e^{-\frac{1}{2}\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}\right\}$$
$$+\sum_{i=n_1+1}^{r}\left\{-\frac{\theta_1}{\theta_2}\left[\frac{t_{i:n}-\tau}{\tau\theta_2+(t_{i:n}-\tau)\theta_1}\right]+\left(\frac{t_{i:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)\left(\frac{t_{i:n}-\tau}{\theta_2^2}\right)\right\}=0. \tag{21}$$

$$\frac{\partial\log l}{\partial y}=\frac{\theta_1}{\tau\theta_2+(y-\tau)\theta_1}$$
$$+\frac{1}{\theta_2}\left\{(s-r-1)\frac{\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)\times e^{-\frac{1}{2}\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}}{e^{-\frac{1}{2}\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}-e^{-\frac{1}{2}\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2}}-(n-s+1)\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)\right\}=0. \tag{22}$$

Since Eq.'s (20)-(22) cannot be solved explicitly, numerical techniques will be used to solve them simultaneously, which leads to find the maximum likelihood predictor (MLP) of Y and the predictive maximum likelihood estimators (PMLEs) of θ_1 and θ_2. The resulting MLP of Y is denoted by $\hat{Y}_M$.

3.2 Conditional median predictor

Raqab and Nagaraja [19] proposed a point predictor based on the conditional distribution of Y given $\boldsymbol{T}=\boldsymbol{t}$, known as conditional median predictor (CMP). A predictor $\hat{Y}$ is called the CMP of Y, if it is the median of the conditional distribution of Y given $\boldsymbol{T}=\boldsymbol{t}$, that is

$$P\left(Y\le\hat{Y}\middle|\boldsymbol{T}=\boldsymbol{t}\right)=P_\theta\left(Y\ge\hat{Y}\middle|\boldsymbol{T}=\boldsymbol{t}\right).$$

Using the conditional distribution of Y given $\boldsymbol{T}=\boldsymbol{t}$, we can obtain

$$P\left(Y\le\hat{Y}\middle|\boldsymbol{T}=\boldsymbol{t}\right)=$$
$$P\left(1-e^{-\frac{1}{2}\left[\left(\frac{Y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]}\ge 1-e^{-\frac{1}{2}\left[\left(\frac{\hat{Y}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]}\middle|\boldsymbol{T}=\boldsymbol{t}\right).$$

It can be shown that, given $\boldsymbol{T}=\boldsymbol{t}$, the distribution of

$$W=1-e^{-\frac{1}{2}\left[\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]},$$

is a Beta distribution with parameters $s-r$ and $n-s+1$, denoted by $Beta\ (s-r,n-s+1)$. So, if B is a random $Beta\ (s-r,n-s+1)$, and M_B represents the median of B, the CMP of Y can be obtained as

$$\hat{Y}_{CMP}=\tau-\frac{\theta_2}{\theta_1}\tau+\theta_2\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log(1-M_B)}. \tag{23}$$

The CMP of Y can be computed approximately by replacing θ_1 and θ_2 in Eq. (23) by their corresponding MLEs.

3.3 Best Unbiased Predictor

A point predictor $\hat{Y}$ of $Y=T_{s:n}$ is called a best unbiased predictor (BUP) of Y, if the mean of its prediction error, $E(\hat{Y}-Y)$ is zero and the variance of its prediction error, $Var(\hat{Y}-Y)$ is less than or equal to that of any other unbiased predictor of Y. Using the conditional density of Y given $\boldsymbol{T}=\boldsymbol{t}$, as in Eq. (15), the BUP of Y is given by

$$\hat{Y}_{BUP}=E(Y|\boldsymbol{T})=\int_{t_{r:n}}^{\infty}y\,g_{T_{s:n}|\boldsymbol{T}}(y|\theta_1,\theta_2,data)\,dy.$$

Using the binomial expansion:

$$\left\{1-e^{-\frac{1}{2}\left[\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]}\right\}^{s-r-1}=$$
$$\sum_{k=0}^{s-r-1}\binom{s-r-1}{k}(-1)^{s-r-k-1}$$
$$\times e^{-\frac{1}{2}(s-r-k-1)\left[\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]},$$

we obtain

$$\hat{Y}_{BUP}=\frac{(s-r)}{\theta_2}\binom{n-r}{s-r}$$

$$\times \sum_{k=0}^{s-r-1} \left\{ \binom{s-r-1}{k} (-1)^{s-r-k-1} \, e^{\frac{1}{2}(n-r-k)\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2} \times \int_{t_{r:n}}^{\infty} \left(\frac{y^2-\tau}{\theta_2}+\frac{\tau y}{\theta_1}\right) e^{-\frac{1}{2}(n-r-k)\left(\frac{y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2} dy \right\}. \quad (24)$$

The BUP of Y can be approximated by substituting the MLEs of the unknown parameters θ_1 and θ_2 in Eq. (24).

4 Prediction intervals

Another aspect of prediction problem is to predict the future unobserved lifetimes by constructing prediction intervals (PIs) for $Y = T_{s:n}, s = r+1, \dots, n$ based on the Type-II censored sample $\boldsymbol{T} = (T_{1:n}, T_{2:n}, \dots, T_{r:n})$. The pivotal, highest conditional density, and shortest-length based methods are considered in this section.

4.1 Pivotal-based PIs

Let us consider the random variable

$$W = 1 - e^{-\frac{1}{2}\left[\left(\frac{Y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2 - \left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]}, Y > t_r. \quad (25)$$

Since the distribution of W given $\boldsymbol{T} = \boldsymbol{t}$, is a Beta distribution with parameters $s-r$ and $n-s+1$, then W can be considered as a pivotal quantity for obtaining the PI of Y. By considering $(1-\alpha), 0 < \alpha < 1$, as a prediction coefficient and using Eq. (25), we obtain

$$P\left(B_{\frac{\alpha}{2}} < W < B_{1-\frac{\alpha}{2}}\right) = 1 - \alpha,$$

where B_α is the $100\,\alpha^{th}$ percentile of the distribution $Beta\,(s-r, n-s+1)$. Therefore, a $(1-\alpha)100\%$ pivotal PI of Y is $(L_1(\boldsymbol{T}), U_1(\boldsymbol{T}))$, where

$$L_1(\boldsymbol{T}) = \tau - \frac{\theta_2}{\theta_1}\tau + \theta_2 \sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2 - 2\log\left(1 - B_{\frac{\alpha}{2}}\right)}$$

$$U_1(\boldsymbol{T}) = \tau - \frac{\theta_2}{\theta_1}\tau + \theta_2 \sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2 - 2\log(1 - B_{1-\frac{\alpha}{2}})}.$$

The prediction limits $L_1(\boldsymbol{T})$ and $U_1(\boldsymbol{T})$ can be evaluated approximately by replacing θ_1 and θ_2 by their corresponding MLEs.

4.2 Highest conditional density PIs

As described above, the conditional distribution of

$$W = 1 - e^{-\frac{1}{2}\left[\left(\frac{Y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2 - \left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]}$$

given $\boldsymbol{T} = \boldsymbol{t}$ is $Beta\,(s-r, n-s+1)$. Therefore, the conditional pdf of W is:

$$g(w) = \frac{(n-r)!}{(s-r-1)!\,(n-s)!} w^{s-r-1}(1-w)^{n-s}, \quad 0 < w < 1. \quad (26)$$

The density in Eq. (26) is unimodal function in w. An interval (d_1, d_2) is called highest conditional density (HCD) PI of content $(1-\alpha)$ if $(d_1, d_2) = \{d : d \in [0,1], f(d) \geq k\} \subseteq [0,1]$, where

$$\int_{d_1}^{d_2} g(w)\,dw = 1 - \alpha,$$

for some $k > 0$. Now, if $r+1 < s < n$, then $g(w)$ is a unimodal function, and it attains its maximum value at $\delta = \frac{s-r-1}{n-r-1} \in (0,1)$. Following Theorem 9.3.2 of Casella and Berger [9], the HCD PI can be obtained by finding two percentiles d_1 and d_2 such that $P(W < d_1) = P(W > d_2) = \frac{\alpha}{2}$, with $d_1 \leq \delta \leq d_2$, satisfying

$$\int_{d_1}^{d_2} g(w)\,dw = 1 - \alpha, \quad (27)$$

and,

$$g(d_1) = g(d_2). \quad (28)$$

Eq.'s (27) and (28) can be simplified as

$$B_{d_2}(s-r, n-s+1) - B_{d_1}(s-r, n-s+1) = 1 - \alpha, \quad (29)$$

and

$$\left(\frac{1-d_2}{1-d_1}\right)^{n-s} = \left(\frac{d_1}{d_2}\right)^{s-r-1}, \quad (30)$$

where

$$B_v(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^v u^{a-1}(1-u)^{b-1}du,$$

is the incomplete beta function and $\Gamma(.)$ is the gamma function. Consequently, a $(1-\alpha)100\%$ HCD PI of Y is given by $(L_2(\boldsymbol{T}), U_2(\boldsymbol{T}))$, with

$L_2(\boldsymbol{T}) =$

$$\tau-\frac{\theta_2}{\theta_1}\tau+\theta_2\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log(1-d_1)},$$

$U_2(\boldsymbol{T})=$

$$\tau-\frac{\theta_2}{\theta_1}\tau+\theta_2\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log(1-d_2)}$$

For the special case when $s=r+1$ and $s<n$, $g(w)=(n-r)(1-w)^{n-r-1}, 0<w<1$, which is a decreasing function in w with $g(0)=n-r$ and $g(1)=0$. Therefore, the PI for Y is of the form $(0,\ d_2)$ such that $d_2=1-\alpha^{1/(n-r)}$. This concludes that

$$L_2(\boldsymbol{T})=t_{r:n},$$

$$U_2(\boldsymbol{T})=\tau-\frac{\theta_2}{\theta_1}\tau+\theta_2\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-\frac{2}{n-r}\log(\alpha)}.$$

When $s=r+1$ and $s=n, g(w)$ is uniform $U(0,1)$. Here d_1 and d_2 are taken such that $d_1=\alpha/2$ and $d_2=1-\alpha/2$. So, we have

$$L_2(\boldsymbol{T})=\tau-\frac{\theta_2}{\theta_1}\tau+\theta_2\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log\left(1-\frac{\alpha}{2}\right)},$$

and

$$U_2(\boldsymbol{T})=\tau-\frac{\theta_2}{\theta_1}\tau+\theta_2\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log\left(\frac{\alpha}{2}\right)}.$$

Finally, when $s=n$ and $s>r+1$, the density $g(w)=(n-r)w^{n-r-1}, 0<w<1$, is increasing function with $g(0)=0$ and $g(1)=n-r$. In this case, we select the PI for Y to be of the form $(d_1,1)$ such that $\int_{d_1}^{1}g(w)dw=1-\alpha$, which implies that $d_1=\alpha^{\frac{1}{n-r}}$. So, a $(1-\alpha)100\%$ HCD PI of Y is given by

$$L_2(\boldsymbol{T})=\tau-\frac{\theta_2}{\theta_1}\tau+\theta_2\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log\left(1-\alpha^{\frac{1}{n-r}}\right)},$$

and $U_2(\boldsymbol{T})=\infty$.

4.3 Shortest-Length based Method

Based on the fact that the conditional distribution of $W=1-e^{-\frac{1}{2}\left[\left(\frac{Y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]}$ given $\boldsymbol{T}=\boldsymbol{t}$ is a $Beta\ (s-r,n-s+1)$, we select the constants c and d that satisfy the equation:

$$P\left(c<1-e^{-\frac{1}{2}\left[\left(\frac{Y-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2\right]}<d\right)=1-\alpha.$$

Here, the constants c and d are chosen to minimize the length of the PI $U_3(\boldsymbol{T})-L_3(\boldsymbol{T})$.
The optimization problem for figuring out the shortest-length (SL) $(1-\alpha)100\%$ PI can be expressed as:
Minimize Length $=U_3(\boldsymbol{T})-L_3(\boldsymbol{T})$
Subject to

$$B_d(s-r,n-s+1)-B_c(s-r,n-s+1)=1-\alpha.$$

The SL $(1-\alpha)100\%$ PI can be constructed by minimizing the Lagrangian function:

$$R(c,d,\lambda)=\theta_2\left[\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log(1-d)}-\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log(1-c)}\right]-\lambda[\{B_d(s-r,n-s+1)-B_c(s-r,n-s+1)\}-(1-\alpha)],$$

where λ is the Lagrange multiplier. By differentiating R with respect to c, d and λ, respectively, we have:

$$\frac{\partial R}{\partial c}=\frac{-\theta_2}{(1-c)\times\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log(1-c)}}+\lambda p(c,s-r,n-s+1)=0.$$

$$\frac{\partial R}{\partial d}=\frac{\theta_2}{(1-d)\times\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log(1-d)}}-\lambda p(c,s-r,n-s+1)=0.$$

$$\frac{\partial R}{\partial\lambda}=[\{B_d(s-r,n-s+1)-B_c(s-r,n-s+1)\}-(1-\alpha)]=0,$$

where $p(x,a,b)$ represents the density of the distribution $Beta(a,b)$. The above equations can be formed equivalently as:

$$\sqrt{\frac{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log(1-d)}{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log(1-c)}}=\frac{(1-c)p(d,s-r,n-s+1)}{(1-d)p(c,s-r,n-s+1)},\quad(31)$$

and,

$$B_d(s-r,n-s+1)-B_c(s-r,n-s+1)=1-\alpha. \quad (32)$$

The constants c and d are obtained by solving (31) and (32) numerically. Hence, a $(1-\alpha)100\%$ PI of Y, based on this technique, is given by $\left(L_3(\boldsymbol{T}),U_3(\boldsymbol{T})\right)$, such that:

$$L_3(\boldsymbol{T})=\tau-\frac{\theta_2}{\theta_1}\tau+\theta_2\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log(1-c)},$$

$$U_3(\boldsymbol{T})=\tau-\frac{\theta_2}{\theta_1}\tau+\theta_2\sqrt{\left(\frac{t_{r:n}-\tau}{\theta_2}+\frac{\tau}{\theta_1}\right)^2-2\log(1-d)}.$$

5 Simulation study and data analysis

In this section, we conduct a simulation study for computations of the prediction methods that are presented in the above sections and comparing their performance. A real data set is considered to illustrate the different techniques suggested in this paper.

5.1 Simulation study

In this subsection, we perform an intensive Monte Carlo simulation study for performance evaluation of the suggested predictors; which were presented in Section 3. The performances are measured in terms of the biases and the mean square prediction errors (MSPEs) of the predictors. The bias and MSPE of a predictor $\hat{Y}$ of $Y=T_{s:n}(s\geq r+1)$, are defined as

$$bias(\hat{Y})=\frac{1}{M}\sum_{k=1}^{M}(\hat{Y}_k-Y),$$

and,

$$MSPE(\hat{Y})=\frac{1}{M}\sum_{k=1}^{M}(\hat{Y}_k-Y)^2,$$

respectively. In addition, we compare the PIs, that are discussed in Section 4, in terms of their estimated average lengths (ALs) and coverage probabilities (CPs).

Consequently, a Monte Carlo simulation is conducted based on different censoring schemes and sample sizes from the Rayleigh distribution under CEM. For particular values of n,r and s, we generate Type-II censored samples as described in Section 2 based on the following schemes:

Scheme 1: $\theta_1=0.3,\theta_2=0.8$, and $\tau=0.25$.

Scheme 2: $\theta_1=1,\theta_2=1.5$, and $\tau=0.9$.

In both cases, we find the value of the point predictors; MLP, CMP and BUP. Moreover, we compute 95% PIs based on pivotal quantity, HCD and SL methods. Type-II censored samples from Rayleigh distribution were randomly generated under these two different schemes with 2000 replications of the simulation process. Using these random samples, prediction biases and MSPEs of the predictors are computed. The so obtained results are presented in Table 1. In Table 2, we have presented the ALs and CPs of the PIs.

Based on these tables, we observe the following remarks

1. For fixed values of n and r, biases and the MSPEs of the point predictors increase as s increases, which is due to the variation of the lifetime to be predicted as s gets large.

2. The prediction biases of the BUP are smaller than those of the CMP and the MLP for all the considered cases. However, the biases resulted from the CMP are close to the biases of the BUP and smaller than the biases of the MLP. By considering the MSPE as an optimality criterion, it is observed that, the CMP outperforms both the BUP and the MLP. Further, it is noticed that the MSPEs of the three predictors are close to each other's, especially when s is close to r, which can be explained by observing the closeness of the MLEs of the parameters and the corresponding PMLEs in most of the considered cases.

3. The PIs obtained using the SL method is more efficient than other methods based on the AL criterion. Its performance tends to be higher when s gets large. On the other hand, it is observed that the HCD PIs outperforms the pivot PIs in the sense of ALs when s tends to be close to r. As s approaches n, the pivot PIs are competeive. Based on the CP criterion, the HCD PIs are superior to the PIs obtained by SL and pivot methods. The CPs of SL and pivot PIs are very close. It is evident that the CPs of all obtained PIs increase when s increases. In this sense, the worse CP occurs when the lifetime to be predicted is immediately after the last observed lifetime.

As a summary, for point prediction aspect, the CMP is the best predictor as it is computationally attractive and has good performances in terms of the biases and MSPE criteria. For prediction interval part, the SL method produces efficient PIs over other methods based on the AL and CP criteria.

Table 1: Biases and MSPEs of the point predictors for the censored lifetimes

Scheme 1: $\theta_1 = 0.3, \theta_2 = 0.8$, and $\tau = 0.25$.							
(n, r)	s	MLP		CMP		BUP	
		Bias	MSPE	Bias	MSPE	Bias	MSPE
(30, 20)	22	-0.0615	0.0245	-0.0218	0.0233	-0.0071	0.0237
	24	-0.0673	0.0326	-0.0175	0.0315	-0.0033	0.0323
	26	-0.1041	0.0509	-0.0420	0.0473	-0.0279	0.0478
	28	-0.1177	0.0682	-0.0344	0.0640	-0.0174	0.0651
	30	-0.2057	0.1478	-0.0501	0.1360	-0.0106	0.1411
(40, 25)	27	-0.0433	0.0167	-0.0157	0.0161	-0.0050	0.0164
	30	-0.0560	0.0242	-0.0201	0.0233	-0.0101	0.0236
	35	-0.0825	0.0469	-0.0275	0.0456	-0.0174	0.0461
	38	-0.1187	0.0829	-0.0381	0.0808	-0.0237	0.0820
	40	-0.1924	0.1502	-0.0416	0.1465	-0.0038	0.1523
(50, 30)	32	-0.0294	0.0123	-0.0080	0.0121	0.0005	0.0124
	35	-0.0413	0.0153	-0.0151	0.0148	-0.0073	0.0150
	40	-0.0568	0.0248	-0.0189	0.0238	-0.0115	0.0240
	45	-0.0819	0.0472	-0.0270	0.0459	-0.0184	0.0463
	50	-0.1971	0.1443	-0.0574	0.1332	-0.0211	0.1365
Scheme 2: $\theta_1 = 1, \theta_2 = 1.5$, and $\tau = 0.9$.							
(n, r)	s	MLP		CMP		BUP	
		Bias	MSPE	Bias	MSPE	Bias	MSPE
(30, 20)	22	-0.1077	0.0878	-0.0320	0.0843	-0.0044	0.0866
	24	-0.1203	0.1143	-0.0244	0.1135	0.0018	0.1173
	26	-0.1927	0.1760	-0.0702	0.1646	-0.0439	0.1671
	28	-0.2593	0.2920	-0.0919	0.2761	-0.0605	0.2803
	30	-0.4489	0.6233	-0.1366	0.5729	-0.0635	0.5903
(40, 25)	27	-0.0930	0.0635	-0.0408	0.0607	-0.0210	0.0613
	30	-0.0959	0.0871	-0.0249	0.0856	-0.0062	0.0874
	35	-0.1524	0.1682	-0.0394	0.1675	-0.0204	0.1703
	38	-0.2459	0.3144	-0.0790	0.3043	-0.0522	0.3086
	40	-0.4218	0.6127	-0.1264	0.5870	-0.0571	0.6042
(50, 30)	32	-0.0634	0.0463	-0.0232	0.0453	-0.0074	0.0459
	35	-0.0807	0.0587	-0.0282	0.0563	-0.0134	0.0571
	40	-0.1070	0.0986	-0.0289	0.0958	-0.0150	0.0971
	45	-0.1756	0.1728	-0.0613	0.1659	-0.0455	0.1673
	50	-0.4225	0.5844	-0.1366	0.5251	-0.0686	0.5354

Table 2: ALs and CPs of 95% PIs of the censored lifetimes

Scheme 1: $\theta_1 = 0.3, \theta_2 = 0.8$, and $\tau = 0.25$.							
(n, r)	s	Pivotal Method		HCD Method		SL Method	
		AL	CP	AL	CP	AL	CP
(30, 20)	22	0.3579	0.654	0.3289	0.653	0.3288	0.656
	24	0.5233	0.831	0.5079	0.831	0.5034	0.832
	26	0.7059	0.883	0.7142	0.885	0.6899	0.881
	28	0.9489	0.918	1.0519	0.933	0.9323	0.907
	30	1.7211	0.951	∞	0.966	1.6782	0.942
(40, 25)	27	0.2517	0.593	0.2306	0.574	0.2305	0.575
	30	0.4062	0.806	0.3954	0.805	0.3937	0.809
	35	0.6763	0.892	0.6918	0.886	0.6678	0.891
	38	0.9532	0.914	1.0842	0.935	0.9416	0.901
	40	1.6706	0.955	∞	0.960	1.6323	0.945
(50, 30)	32	0.2046	0.594	0.1871	0.569	0.1871	0.569
	35	0.3209	0.761	0.3111	0.760	0.3105	0.763
	40	0.4776	0.848	0.4764	0.847	0.4715	0.850
	45	0.6874	0.884	0.7123	0.887	0.6816	0.883
	50	1.6841	0.954	∞	0.984	1.6464	0.952
Scheme 2: $\theta_1 = 1, \theta_2 = 1.5$, and $\tau = 0.9$.							
(n, r)	s	Pivotal Method		HCD Method		SL Method	
		AL	CP	AL	CP	AL	CP
(30, 20)	22	0.6719	0.661	0.6175	0.640	0.6173	0.640
	24	0.9939	0.802	0.9646	0.799	0.9560	0.797
	26	1.3014	0.860	1.3167	0.866	1.2719	0.852
	28	1.8087	0.914	2.0051	0.928	1.777	0.907
	30	3.2288	0.952	∞	0.962	3.1485	0.945
(40, 25)	27	0.4768	0.607	0.4368	0.578	0.4367	0.579
	30	0.7790	0.777	0.7583	0.780	0.7551	0.786
	35	1.2733	0.876	1.3026	0.886	1.2574	0.867
	38	1.7843	0.879	2.0295	0.916	1.7626	0.870
	40	3.1547	0.937	∞	0.953	3.0825	0.931
(50, 30)	32	0.3734	0.600	0.3414	0.563	0.3414	0.564
	35	0.6075	0.756	0.5888	0.760	0.5878	0.762
	40	0.9015	0.828	0.8993	0.827	0.8900	0.831
	45	1.2881	0.850	1.3349	0.863	1.2773	0.851
	50	3.1285	0.948	∞	0.952	3.0586	0.945

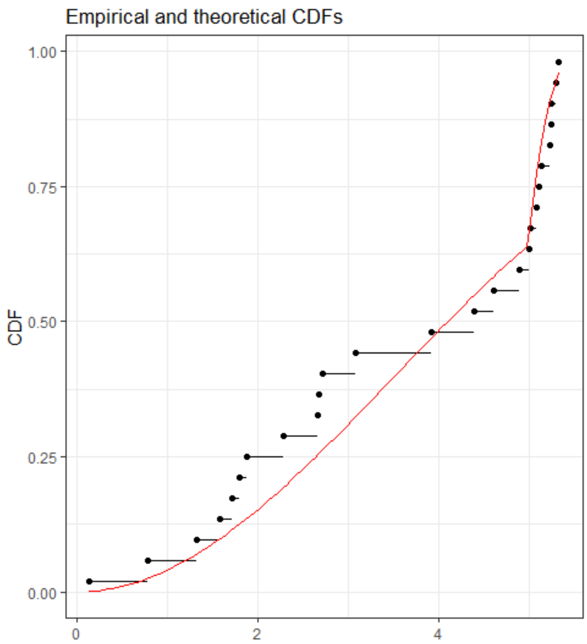

Fig. 1: The empirical cdf (dots); and the estimated cdf of Rayleigh CEM model based on MLE (solid line).

5.2 Data analysis

To clarify the prediction methods proposed in this paper, we perform a real data analysis. The dataset is taken from Han and Kundu [12], it represents a total of 31 failure times (in hundred hours) from a sample of 35 prototypes of a solar lighting device with two dominant failure modes, controller failure and capacitor failure. Here, temperature is the stress factor whose level was changed during the test in the range of 293K to 353K with the normal operating temperature at 293K, and stress change time at 500 hours. These data have been used previously by Kotb and El-Din [13]. The data are recorded in Table 3.

Table 3: Lifetimes of prototypes of a solar lighting device on a simple step-stress test

Temperature Level	Recorded data		
S_1: 293K	0.140	0.783	1.324
	1.582	1.716	1.794
	1.883	2.293	2.660
	2.674	2.725	3.085
	3.924	4.396	4.612
	4.892		
S_2: 353K	5.002	5.022	5.082
	5.112	5.147	5.238
	5.244	5.247	5.305
	5.337	5.407	5.408
	5.445	5.483	5.717

To visualize the accuracy of our model, i.e., Rayleigh CEM, the true cdf of the lifetimes is plotted in Fig. 1, along with the corresponding cdf based on the maximum likelihood estimate. However, it was shown by Kotb and El-Din [15] that Rayleigh distribution is appropriate for analyzing this data set.

Suppose the life test is terminated when the 26^{th} lifetime is observed, i.e., we observe a Type-II censored sample with $n = 35$, $r = 26$. Our purpose is to obtain point predictors of the unobserved lifetimes $Y = T_{s:n}, s = 28, 30, 31, 33, 35$, and the associated PIs.

First we compute the MLEs of θ_1 and θ_2 by solving Eq.'s (12) and (13) simultaneously, it is found that $\hat{\theta}_1 = 4.360$ and $\hat{\theta}_2 = 0.653$. For predicting the future censored lifetimes, point predictors and PIs are reported in Table 4.

Table 4: Point predictors and PIs for future lifetimes of $Y = T_{s:n}$.

Point predictors of $Y = T_{s:n}$				
s	True value	MLP	CMP	BUP
28	5.408	5.379	5.412	5.425
30	5.483	5.475	5.518	5.531
31	5.717	5.533	5.582	5.594
33	-----	5.682	5.750	5.765
35	-----	5.967	6.095	6.128
95% PIs of $Y = T_{s:n}$				
s	True value	Pivotal PI	HCD PI	SL PI
28	5.408	(5.348, 5.571)	(5.340, 5.541)	(5.339, 5.540)
30	5.483	(5.393, 5.737)	(5.388, 5.723)	(5.377, 5.707)
31	5.717	(5.426, 5.833)	(5.427, 5.832)	(5.409, 5.803)
33	-----	(5.522, 6.093)	(5.540, 6.166)	(5.500, 6.057)
35	-----	(5.703, 6.740)	(5.753, ∞)	(5.657, 6.663)

It can be observed that the point predictors are close to the true values, with advantage to the CMP. Moreover, the point predictors obtained are lying within all considered PIs. It can be observed that all PIs obtained contain the true values of the future order statistics. It is also noticed that the PIs become wider when s gets large, the reason is that the fluctuation of $Y = T_{s:n}$ tends to be high as Y moves away from the observed failures times. Although all PIs are close in the sense of AL criterion, the PIs constructed by SL method have shortest lengths.

6 Conclusions

In this paper, we have addressed the prediction of future lifetimes of a simple step stress test of Rayleigh distribution under CEM when the data are Type-II censored. Several point predictors are proposed including, maximum likelihood, conditional median, and best unbiased predictors. We have also discussed another aspect of prediction, which is constructing prediction intervals for the future lifetimes. We have compared the performance of the predictors obtained by extensive Monte Carlo simulation study by considering the biases and MSPEs of the suggested predictors. Prediction intervals were also assessed in terms of the average lengths and coverage probabilities. It is observed that the CMP has the best performance among all point predictors. In the context of interval prediction, it is observed that the SL based method is the most suitable method for obtaining PIs of future lifetimes.
It is worth mentioning that the results of this paper were mainly obtained for Type-II censored scheme, but our techniques can be performed for other censoring schemes, as Type-I, hybrid or progressive censoring.

References

[1]. Ababneh, O. Y., New Newton's method with third-order convergence for solving nonlinear equations, *World Academy of Science, Engineering and Technology*, 6, 2012, 118-120.

[2]. Amleh, M. A., and Raqab, M. Z., Inference in simple step-stress accelerated life tests for Type-II censoring Lomax data. *Journal of Statistical Theory and Applications*, 20(2), 2021, 364-379.

[3]. Amleh, M. A., and Raqab, M. Z., Prediction of censored Weibull lifetimes in a simple step-stress plan with Khamis-Higgins model, *Statistics, Optimization and Information Computing*, 2021, doi: 10.19139/soic-2310-5070-1069

[4]. Bai, D. S., and Kim, M. S., Optimum simple step- stress accelerated life tests for Weibull distribution and Type I censoring, *Naval Research Logistics*, 40(2), 1993, 193-210.

[5]. Balakrishnan, N., Zhang, L., and Xie, Q., Inference for a simple step-stress model with Type-I censoring and lognormally distributed lifetimes, *Communications in Statistics- Theory and Methods*, 38(10), 2009, 1690-1709.

[6]. Basak, I., Prediction of times to failure of censored items for a simple step-stress model with regular and progressive Type-I censoring from the exponential distribution, *Communications in Statistics- Theory and Methods*, 43(10), 2014, 2322-2341.

[7]. Basak, I., and Balakrishnan, N., Prediction of censored exponential lifetimes in a simple step-stress model under progressive Type-II censoring, *Computational Statistics*, 32(4), 2016, 1665-1687.

[8]. Basak, I., and Balakrishnan, N., A note on the prediction of censored exponential lifetimes in a simple step-stress model with Type-II censoring. *Calcutta Statistical Association Bulletin*, 70(1), 2018, 57-73.

[9]. Casella, G., and Berger, R. L., *Statistical inference*, 2nd edition, Pacific Grove, CA, Duxbury, 2002.

[10]. Chandra, N., and Khan, M. A., Estimation of parameters in step-stress accelerated life tests for the Rayleigh distribution under censoring setup, *Journal of Reliability and Statistical Studies*, 7(2), 2014, 1-18.

[11]. Ebrahem, M. A. H., and Al-Masri, A. Q., Estimating the parameters of Rayleigh cumulative exposure model in simple step-stress testing, *Journal of Modern Applied Statistical Methods*, 8(2), 2009, 478-487.

[12]. Han, D. and Kundu, D., Inference for a step-stress model with competing risks for failure from the generalized exponential distribution under Type-I censoring, *IEEE Transactions on Reliability*, 64(1), 2015, 31–43.

[13]. Johnson, N.L., Kotz, S., and Balakrishnan, N., *Continuous Univariate Distribution (Volume 1),* John Wiley & Sons, New York, 1994.

[14]. Kaminsky, K. S., and Rhodin, L. S., Maximum likelihood prediction, *Annals of the Institute of Statistical Mathematics*, *37*(3), 1985, 507-517.

[15]. Kotb, M. S. and El-Din, M. M., Parametric Inference for step-stress accelerated life testing from Rayleigh distribution under ordered ranked set sampling, *IEEE Transactions on Reliability*, 2020, doi: 10.1109/TR.2020.2999392

[16]. Kumar, M., Pathak, A., and Soni, S., Bayesian inference for Rayleigh distribution under step-stress partially accelerated test with progressive Type-II censoring with binomial removal, *Annals of Data Science*, 6(1), 2019, 117-152.

[17]. Kundu, D. and Ganguly, A., *Analysis of Step-Stress Models: Existing Results and Some Recent Developments*, Academic Press, London, 2017.

[18]. Mitra, S., Ganguly, A., Samanta, D., and Kundu, D., On the simple step-stress model for two-parameter exponential distribution, *Statistical Methodology*, 15, 2013, 95-114.

[19]. Nelson, W., Accelerated life testing-step-stress models and data analyses, *IEEE transactions on reliability*, 29(2), 1980, 103-108.

[20]. Nelson W.B., *Accelerated Testing: Statistical Models, Test Plans, Data Analyses*, John Wiley & Sons, New York, 1990.

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

[22]. Xiong C., Inferences on a simple step-stress model with Type II censored exponential data, *IEEE Transactions on Reliability*, 55, 1998, 67-74.