An Efficient Estimator of Reliability for Exponential Class Software Reliability Models

B. Roopashri Tantri and Murulidhar N. N.

Abstract—Software quality has become a major concern of all software manufacturers. One such measure of software quality is the reliability, which is the probability of failure-free operation of a software in a specified environment for a specified time. If \( T \) denotes the time to failure of any software, then, the reliability of this software, denoted by \( R(t) \), is given by \( R(t)=P(T>t) \). The reliability of any software can be estimated using various methods of estimation. The simplest among these methods, is the method of maximum likelihood estimation. Even though it satisfies most of the desirable properties of a good estimator, it is still not as efficient as the minimum variance unbiased estimator. In this paper, the minimum variance unbiased estimator of \( R(t) \) for exponential class software reliability models, is obtained using a procedure called blackwellization. The estimator so obtained using this method always has minimum variance. The same is verified for a model belonging to the exponential class, viz, the Jelinski - Moranda model.

I. INTRODUCTION

The increasing pace of change in computing technology means that information systems become economically obsolete more rapidly. Software can become obsolete either as a result of new hardware or software technology. The most important software product characteristics are quality, cost and schedule. Quantitative measures exist for the latter two characteristics, but the quantification of quality has become more difficult. It is most important, however, because the absence of a concrete measure for software quality generally makes that quality will suffer when it competes for attention against cost and schedule. One of the important measures of software quality is the reliability. Software reliability concerns itself with how well the software functions to meet the requirements of the customer. It is the probability that the software will work without failure for a specified period of time [1]. It relates to operation rather than design of the program. In general, reliability is the probability of failure free operation of software in a specified environment for a specified period of time [2]. The failure times are random in nature. If \( T \) denotes the time to failure of software, then, its reliability is given by \( R(t) = P(T > t) \). For any given software, this reliability can be estimated using various methods of estimation. The most popular method is the method of maximum likelihood estimation (MLE). In case of exponential class models, where failure time is assumed to be exponential, the MLE of \( R(t) \) is easy to obtain and hence is widely used [3]. But, usually such an MLE is biased and is not efficient. To obtain an unbiased and more efficient estimator of \( R(t) \), the method of minimum variance unbiased estimation (MVUE) is used.

Some terminologies used:

- **Software reliability** [1]: It is the probability of failure-free operation of a software in a specified environment for a specified time.
- **Software reliability models** [1], [4]: These describe the behavior of failure with time, by expressing failures as random processes in either times of failure or the number of failures, at fixed times.
- **Exponential class software reliability models** [2], [4]: This group consists of all finite failure models which have exponential failure time distribution.
- **Estimator**: A function of the sample observations used to estimate the value of the parameter of any distribution is said to be an estimator.
- **Any estimator is expected to possess certain statistical properties. An estimator that satisfies most of such desirable properties is considered as the best estimator.**

The statistical properties of good estimators are given below:

- **Statistical properties of good estimators** [3]: A good estimator should be 1) Consistent 2) Unbiased 3) Sufficient 4) Efficient.

1. **Consistency**: An estimator \( T \) based on a sample of size \( n \) from a distribution with parameter \( \theta \) is said to be consistent for \( \theta \), if \( T \) converges in probability to \( \theta \), i.e., if \( P(|T-\theta|<\varepsilon )\to 1 \) as \( n \to \infty \), for every \( \varepsilon >0 \).

2. **Unbiasedness**: An estimator \( T \) based on a sample of size \( n \) from a distribution with parameter \( \theta \), is said to be unbiased for \( \theta \), if \( E(T)=\theta \).

3. **Sufficiency**: An estimator \( T \) is said to be a sufficient estimator for the parameter \( \theta \), if it contains all the information in the sample regarding that parameter. Or, in other words, an estimator \( T \) is sufficient for the parameter \( \theta \) if the conditional distribution of the sample given \( T \) is independent of \( \theta \). Using the factorization theorem, a statistic \( T \) is sufficient for the parameter \( \theta \), if the likelihood function \( L \) can be expressed as \( L=g(T, \theta)h(x) \), where in \( g, \theta \) depends on \( x \) only through \( T \) and \( h(x) \) is independent of \( \theta \).

4. **Efficiency**: If in a class of estimators, there exists one whose variance is less than that of any such estimator, then it is called an efficient estimator. Thus, if \( T_1 \) and \( T_2 \) are two estimators, then \( T_1 \) is more efficient than \( T_2 \), if
\( V(T_1) < V(T_2) \).

**Complete estimator:** Let \( T \) be an estimator and let \( h(T) \) be any function of \( T \). Then, \( T \) is said to be the complete estimator, if \( E[h(T)] = 0 \) implies \( h(T) = 0 \).

**Complete sufficient estimator:** A sufficient estimator which is also complete is called the complete sufficient estimator.

II. METHODS OF ESTIMATION

In order to estimate the reliability, various methods of estimation are available. Some of these methods are – Method of maximum likelihood estimation (MLE), Method of least squares, Method of moments, Method of minimum variance unbiased estimation (MVUE) etc. All these estimators should satisfy the statistical properties described above. The one that satisfies most of these statistical properties is considered as the best estimator.

Two methods of estimation viz, the method of MLE and the method of MVUE are explained below.

A. Method of Maximum Likelihood Estimation (MLE) [3]

Let \( x_1, x_2, x_3, \ldots , x_n \) be a random sample of size \( n \) from a population with probability function \( f(x, \theta) \). Then, the likelihood function is given by

\[ L = \prod_{i=1}^{n} f(x_i, \theta) \]

We have to find an estimator that maximizes the likelihood function.

By differential calculus, this means

\[ \frac{\partial L}{\partial \theta} = 0 \quad \text{with} \quad \frac{\partial^2 L}{\partial \theta^2} < 0 \]

Since \( L > 0 \), we can write the above as,

\[ \frac{1}{L} \frac{\partial L}{\partial \theta} = 0 \quad \text{i.e.,} \quad \frac{\partial (\ln L)}{\partial \theta} = 0. \]

Thus, MLE is the solution of

\[ \frac{\partial (\ln L)}{\partial \theta} = 0 \quad \text{with} \quad \frac{\partial^2 (\ln L)}{\partial \theta^2} < 0 \]

B. Method of Minimum Variance Unbiased Estimation (MVUE) [4]

If a statistic \( T \) based on a sample of size \( n \) is such that - \( T \) is unbiased for \( \theta \) and has the smallest variance among the class of all unbiased estimators of \( \theta \), then, \( T \) is called the MVUE of \( \theta \). It is found that such an MVUE is always unique. To find this MVUE, we start with an unbiased estimator \( U \) and then improve upon it by defining a function \( \Phi(T) \) of the sufficient statistic \( T \). This procedure, called Blackwellisation is explained below:

Let \( U = U(x_1, x_2, \ldots , x_n) \) be an unbiased estimator of parameter \( \theta \) and let \( T = T(x_1, x_2, \ldots , x_n) \) be a sufficient statistic for \( \theta \). Consider the function \( \Phi(T) \) of this sufficient statistic, defined as \( \Phi(T) = E[U|T=t] \). Then, \( E[\Phi(T)] = 0 \) and \( V[\Phi(T)] \leq V(U) \). If in addition, \( T \) is also complete, then \( \Phi(T) \) is also the unique estimator.

The MVUE as obtained using the method of Blackwellisation always gives an estimator that is more efficient than any other estimator. That means it always gives an estimator that has less variance than any other estimator.

III. RELIABILITY ESTIMATION FOR EXPONENTIAL CLASS MODELS

MLE is a very simple and widely used method of estimation. MLEs are always consistent and sufficient. But they need not be unbiased and efficient. But, MVUEs are always unbiased and efficient. Thus they are considered as the best estimators among the class of all estimators [5].

Let us consider the problem of finding MLE and MVUE of \( R(t) \) and compare their measures of dispersion, viz, the variance, to find the best among the two.

**MLE of \( R(t) \):** The maximum likelihood estimators have a property that, if \( \hat{\theta} \) is an MLE of \( \theta_0 \), and if \( g(\hat{\theta}) \) is any function of \( \hat{\theta} \), then \( g(\hat{\theta}) \) is the MLE of \( g(\theta_0) \). This is called the invariance property. Using this property, we can obtain the MLE of \( R(t) \) as follows:

In the exponential class models, it is assumed that the time to failure of a software is exponentially distributed with parameter \( \Phi \), the failure rate. Thus, if \( T \) denotes the time to failure, then \( T \sim \exp(\Phi) \). Hence, its probability function is given by \( f(t) = \Phi e^{-\Phi t} \). Using the above procedure of finding MLE, we get the MLE of \( \Phi \) as \( \hat{\Phi} = \frac{n}{\sum_{i=1}^{n} t_i} \) [6]. Thus,

\[
\hat{R}(t) = e^{-\hat{\Phi} t} = e^{-\frac{n}{\sum_{i=1}^{n} t_i} t}.
\]

**MVUE of \( R(t) \):** To find the MVUE of \( R(t) = P(T \geq t) \), define a function \( U(t_1, t_2, \ldots , t_m) \), such that

\[
U(t_1) = \begin{cases} 1 & \text{if } t \geq t_i \\ 0 & \text{otherwise} \end{cases}
\]

Thus, \( E[U(t_1)] = 1 \cdot P(T \geq t) + 0 \cdot P(T < t) = P(T \geq t) = R(t) \). This shows that \( U(t_1) \) is unbiased for \( R(t) \).

Also, the likelihood function is given by

\[
L = \prod_{i=1}^{n} f(t_i) = \Phi^n e^{-\Phi \sum_{i=1}^{n} t_i}
\]

Using factorization theorem, \( \sum_{i=1}^{n} t_i \) is the complete sufficient estimator for the parameter \( \Phi \).

Thus, MVUE of \( R(t) \) is obtained as

\[
\hat{R}(t) = E[U(t_1) \sum_{i=1}^{n} t_i] = \int_{t}^{\infty} f(t) \sum_{i=1}^{n} t_i \right dt
\]
but,
\[
f(t|\sum_{i=1}^{n} t_i) = \frac{f(t,\sum_{i=1}^{n} t_i)}{f(\sum_{i=1}^{n} t_i)} \tag{3}
\]

It can be shown that
\[
f(\sum_{i=1}^{n} t_i) = \frac{\Phi^n}{\Gamma(n)} e^{-\Phi \sum_{i=1}^{n} t_i} (\sum_{i=1}^{n} t_i)^{n-1} \tag{4}
\]

To find the probability function \( f(t,\sum_{i=1}^{n} t_i) \), we split the sample \( (T_1, T_2, T_3, \ldots, T_n) \) into two samples as \( T_i = t_i \) of size one and \( (T_2, T_3, \ldots, T_n) \) of size \( n-1 \). Noting that \( T_i \)'s are independent and are exponentially distributed, we get the joint probability function of \( T_1 \) and \( \sum_{i=2}^{n} t_i \) as
\[
f(t,\sum_{i=1}^{n} t_i) = e^{-\Phi \sum_{i=1}^{n} t_i} \Phi^t (\sum_{i=1}^{n} t_i)^{n-1} \tag{5}
\]

Now, using (4) and (5) in (3) and simplifying, we get the conditional probability function as
\[
f(t|\sum_{i=1}^{n} t_i) = \left( n-1 \right) \left( 1 - \frac{t}{\sum_{i=1}^{n} t_i} \right)^{n-1} \text{ if } 0 < t < \sum_{i=1}^{n} t_i \]
0 otherwise

Substituting the above probability function in (2), we get the estimate (MVUE) of reliability as
\[
\hat{R}(t) = \left\{ \begin{array}{ll}
\left( \frac{n-t}{n} \right)^{t} \left( \frac{t}{\sum_{i=1}^{n} t_i} \right)^{n-t} & \text{if } t < \sum_{i=1}^{n} t_i \\
0 & \text{otherwise}
\end{array} \right. \tag{6}
\]

IV. DISCUSSION

The MVUE obtained using the method of Blackwellization is always unbiased and has minimum variance. Let us check this with the help of a case study, by comparing the variance of MLE of \( R(t) \) with that of MVUE of \( R(t) \). Let us consider two case studies which give failure times of different failures observed by Lyu [1] for an exponential class model, viz, the Jelinski - Moranda model [7].

**Case study I:** Table I gives the failure time of 10 failures. For this data, the MLE of the failure rate \( \Phi \) is given by \( \hat{\Phi} = 0.01335 \). The MLE and MVUE of the reliability function \( R(t) \) are given by
\[
\hat{R}(t) = e^{-0.01335t} \quad \text{and} \quad \hat{R}(t) = 0.9 \left( 1 - \frac{t}{749} \right)^{20}
\]

Calculations in Table II show the values of MLE and MVUE of \( R(t) \) and also their sum of squares of errors. Fig. 1 shows the graph of these two curves. In these graphs, time of failure is taken along x axis and \( R(t) \) is taken along y axis. Using the calculations of Table II, we get
\[
V(\hat{R}(t)) = 0.224649241 \quad \text{and} \quad V(\hat{R}(t)) = 0.177196922.
\]

This establishes that MVUE of \( R(t) \) has variance less than that of MLE of \( R(t) \).

**Case study II:** Table III shows failure times of 21 failures. MLE of failure rate \( \Phi \) is given by \( \hat{\Phi} = 0.00868 \). The MLE and MVUE of reliability are given by
\[
\hat{R}(t) = e^{-0.00868t} \quad \text{and} \quad \hat{R}(t) = 0.9524 \left( 1 - \frac{t}{2417.53} \right)^{21}
\]

![Fig. 1. The graph of these two curves.](image)
That squares of errors. Fig. 2 shows the graph of these two curves, where failure time is taken along x-axis and $R(t)$ along y-axis.

Again, we can see from the calculations of the above table that $V(\hat{R}(t)) = 0.165760866$ and $V(\hat{R}(t)) = 0.145622529$. Again it is established that MVUE of $R(t)$ has less variance than MLE of $R(t)$.

### Table III: Failure Times of 21 Failures

<table>
<thead>
<tr>
<th>Failure Number</th>
<th>Time of Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15.7</td>
</tr>
<tr>
<td>2</td>
<td>29.39</td>
</tr>
<tr>
<td>3</td>
<td>41.14</td>
</tr>
<tr>
<td>4</td>
<td>56.47</td>
</tr>
<tr>
<td>5</td>
<td>75.61</td>
</tr>
<tr>
<td>6</td>
<td>98.83</td>
</tr>
<tr>
<td>7</td>
<td>112.42</td>
</tr>
<tr>
<td>8</td>
<td>125.61</td>
</tr>
<tr>
<td>9</td>
<td>129.39</td>
</tr>
<tr>
<td>10</td>
<td>133.45</td>
</tr>
<tr>
<td>11</td>
<td>138.94</td>
</tr>
<tr>
<td>12</td>
<td>141.41</td>
</tr>
<tr>
<td>13</td>
<td>143.67</td>
</tr>
<tr>
<td>14</td>
<td>144.95</td>
</tr>
<tr>
<td>15</td>
<td>145.16</td>
</tr>
<tr>
<td>16</td>
<td>145.16</td>
</tr>
<tr>
<td>17</td>
<td>146.25</td>
</tr>
<tr>
<td>18</td>
<td>146.7</td>
</tr>
<tr>
<td>19</td>
<td>147.26</td>
</tr>
<tr>
<td>20</td>
<td>148.15</td>
</tr>
<tr>
<td>21</td>
<td>152.4</td>
</tr>
</tbody>
</table>

Table IV shows MLE and MVUE of $R(t)$ and their sum of squares of errors. Fig. 2 shows the graph of these two curves, where failure time is taken along x-axis and $R(t)$ along y-axis.

$$\text{MLE of } R(t) = 0.165760866$$

$$\text{MVUE of } R(t) = 0.145622529$$

Again, we can see from the calculations of the above table that variance of MVUE of $R(t)$ is less than variance of MLE of $R(t)$, which is in tune with the theoretical results. Hence, the reliability as given by (6) gives more accurate value of the reliability for any exponential class software reliability models.

## Conclusion

One of the best measures of software quality is the reliability. The maximum likelihood estimate of reliability can very easily be obtained for the exponential class models. But it is not as efficient as MVUE. As seen earlier, MLE is always consistent and sufficient, but need not be unbiased and efficient. But, MVUE is always unbiased and efficient in addition to being consistent and sufficient. From the case studies above, we can see that variance of MVUE of $R(t)$ is less than variance of MLE of $R(t)$.

## Acknowledgment

The authors would like to thank NCET, Bangalore and NITK, Surathkal for their support in preparing this paper.

## References


B. Roopashri Tantri was graduated in science in the year 1991 from M. G. M. College, Udupi, Karnataka, India with specialization in physics, statistics and mathematics. She completed M. Sc. in statistics from Mangalore University, Karnataka, India in the year 1993 and M. Tech. in systems analysis and computer applications from National Institute of Technology, Karnataka, Surathkal, India in the year 2000.

She has a total of about 20 years of teaching experience in engineering colleges. She had spent 11 years teaching B. E. and M.C.A. courses at NMAMIT, Nitte, Karnataka. Currently, she is working as a professor and the head of the Department of Information Science & Engineering at Mangalore University, Karnataka, India. She has attended several national and international conferences and has published research articles in various national and international journals. Her areas of interest are reliability engineering, software reliability, stochastic analysis & operations research.

Pro. Tantri is a life member of ISSTE (Indian Society for Technical Education) and is also a recipient of K. M. Rai gold medal for securing first rank in M. Sc.

Murulidhar N. N. obtained his Ph. D. in reliability engineering from IIT, Bombay, India in 1989.

He joined National Institute of Technology, Karnataka, Surathkal, India, (Formerly KREC, Surathkal) as a faculty. Presently, he is a professor and the head of the Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka, Surathkal. He has attended several national and international conferences and has published research articles in various national and international journals. His areas of interest are reliability engineering, software reliability, stochastic analysis & operations research.

Pro. Murulidhar is a life member of ISPS, RMS, ISTE and member ISRE.