A Study on the Reliability Performance Analysis of Finite Failure NHPP Software Reliability Model Based on Weibull Life Distribution

Tae-Jin, Yang
Namseoul University, Department of Electronic Engineering, Chungnam, 31020, Korea.
*Corresponding author: solomon645@nsu.ac.kr

Abstract:
In this study, the performance of reliability is analyzed by applying the Weibull life distribution to the finite-fault NHPP reliability model. For this, software failure time data was used, parametric estimation was applied to the maximum likelihood estimation method, and nonlinear equations were calculated using the bisection method. As a result, in the analysis of the intensity function, the Inverse-exponential model is an efficient model because the failure occurring rate decreases with the failure time and the mean square error (MSE) is the smallest. In the analysis of the mean value function, all the proposed models showed a slightly overestimated value compared to the true value, but the Inverse-exponential model showed the smallest error value to the true value. As a result of evaluating the software reliability after putting the mission time in the future, the Inverse-exponential model was stable and high together with the Rayleigh model, but the Goel-Okumoto basic model showed a decreasing tendency. In conclusion, we found that the Inverse-exponential model is an efficient model with the best performance among the proposed models. In this study, the reliability performance of the Weibull life distribution model without the existing research case was newly analyzed, and it is expected that it can be used as a basic guideline for the software developers to exploring the optimal software reliability model.


1. INTRODUCTION
Software technology, which is the core of the 4th industrial revolution era, has spread rapidly in various industrial fields, and the need for software development that can process large amount of information quickly and accurately without failures is also increasing. To solve this problem, software developers are still doing a lot of research to search ways to improve software reliability.

After all, for software developers, solving problems to improve software reliability is a very important topic. For this reason, software reliability models using the non-homogeneous Poisson process (NHPP) have been extensively studied to improve software reliability. In particular, many NHPP software reliability models using the intensity function and the mean value function have been proposed to estimate the reliability attributes such as the number of residual failure and the failure rate in a controlled test environment [1]. In relation to the NHPP reliability model, Goel and Okumoto [2] proposed an exponential software reliability model, Huang [3] explained the software reliability attributes using the mean function, Shyur [4] proposed a generalized reliability model using change-point, and Kim [5] analyzed the attributes of software development costs based on the weibull distribution. In addition, Pham and Zhang [6] proposed a new model based on NHPP software reliability with testing coverage, and Voda [7] proposed that various types of lifetime distributions can be explained by the inverse-Rayleigh distribution. Zhang and Wu [8] also proposed a new software reliability cost model based on software failure time.

Therefore, in this study, after applying the Weibull distribution widely used in the reliability field to the finite-fault NHPP model, we analyze the reliability performance of the NHPP Weibull reliability model, and will present the optimal software reliability model through the analysis results.

2. RELATED RESEARCH
2.1 NHPP Software Reliability Model
N(t) is the cumulative number of failures of the software detected up to time t, m(t) is a mean value function, when λ(t) is expressed by an intensity function, the cumulative failure number N(t) follows a Poisson probability density function having a parameter m(t). The software reliability model of Non-Homogeneous Poisson Process (NHPP) is a model that measures the reliability by using the average failure rate function around the number of failures generated per unit time. That is

\[ P(N(t) = n) = \frac{[m(t)]^n \cdot e^{-m(t)}}{n!} \]  

Note that \( n = 0,1,2,\ldots,\infty \).

The mean value function \( m(t) \) and the intensity function \( \lambda(t) \) of the NHPP model are as follows.

\[ m(t) = \int_0^t \lambda(s) \, ds \]  

\[ \frac{dm(t)}{dt} = \lambda(t) \]  

In terms of software reliability, the mean value function represents a software failure occurrence expected value, the intensity function is the failure rate function and means the

1890
failure occurrence rate per defect. Also, the time domain NHPP models are classified into a finite failure that the failure does not occur at the time of repairing the failure, and an infinite failure that the failure occurs at the time of repairing failure. In this study, we will analyze the software reliability performance based on finite failure cases. That is, in the finite-failure NHPP model, if the expected value of the failure that can be found up to time [0, t] is 0, then the mean value function and the intensity function are as follows.

\[ m(t, \theta, b) = \theta F(t) \]  \hspace{1cm} (4)  

\[ \lambda(t, \theta, b) = \theta F(t)' = \theta f(t) \] \hspace{1cm} (5)  

Considering the failure time data up to the \( n \)th and the Eq. 4 and Eq. 5, the likelihood function of the finite-failure NHPP model is derived as follows:

\[ L_{NHPP}(\theta | x) = \left( \prod_{i=1}^{n} \lambda(x_i) \right) \exp[-m(x_n)] \] \hspace{1cm} (6)  

Note that \( x = (x_1, x_2, x_3, \ldots x_n) \).

### 2.2 Finite Failure NHPP: Goel-Okumoto Basic Model

The Goel-Okumoto model is a well-known basic model in software reliability field. Let \( f(t) \) and \( F(t) \) for the Goel-Okumoto model be a probability density function and a cumulative density function, respectively. Assuming that the expected value of the number of failures of the observation point \([0, t]\) is \( \theta \), the finite failure strength function and the mean value function are as follows.

\[ m(t, \theta, b) = \theta F(t) = \theta(1 - e^{-bt}) \] \hspace{1cm} (7)  

\[ \lambda(t, \theta, b) = \theta f(t) = \theta be^{-bt} \] \hspace{1cm} (8)  

Note that \( \theta > 0, b > 0 \).

Considering the failure time data up to the \( n \)th and the Eq. 7 and Eq. 8, the likelihood function of the finite-failure NHPP model is derived as follows:

\[ L_{NHPP}(\theta, b | x) = \left( \prod_{i=1}^{n} \theta be^{-bx_i} \right) \exp[-\theta(1 - e^{-b x_n})] \] \hspace{1cm} (9)  

Note that \( x = (0 \leq x_1 \leq x_2 \leq \ldots \leq x_n) \).

The likelihood function, using the Eq. 9, is simplified to the following log conditional expression.

\[ \ln L_{NHPP}(\theta | x) = n \ln \theta + n ln b - b \sum_{k=1}^{n} x_k - \theta(1 - e^{-bx_n}) \] \hspace{1cm} (10)  

Therefore, the maximum likelihood estimator \( \hat{\theta}_{MLE} \) and \( \hat{\theta}_{MLE} \) satisfying the following the Eq. 11 and Eq. 12 can be estimated by a numerical method.

\[ \frac{p \ln L_{NHPP}(\theta | x)}{\partial \theta} = \frac{n}{\hat{\theta}} - 1 + e^{-\hat{\theta} x_n} = 0 \] \hspace{1cm} (11)  

\[ \frac{p \ln L_{NHPP}(\theta | x)}{\partial b} = \frac{n}{b} - \sum_{i=1}^{n} x_i - \hat{\theta} x_n e^{-\hat{\theta} x_n} = 0 \] \hspace{1cm} (12)  

### 2.3 Finite Failure NHPP: Rayleigh Model

The Weibull lifetime distribution is widely known as a suitable model for life test and reliability measurements. The probability density function and the cumulative distribution function considering the shape parameter(\( \alpha \)) are as follows [9].

\[ f(t) = \frac{t^{\alpha - 1}}{\beta^\alpha} e^{-\frac{t^\alpha}{\beta^\alpha}} \] \hspace{1cm} (13)  

\[ F(t) = \left( 1 - e^{-\frac{t^\alpha}{\beta^\alpha}} \right) \] \hspace{1cm} (14)  

Note that \( \beta > 0, t \in [0, \infty] \).

In order to simplify the Eq. 13 and Eq. 14, if substitution by the equation \( \frac{1}{\beta^\alpha} = b \) is as follows.

\[ f(t) = 2bt^{\alpha - 1}e^{-bt^\alpha} \] \hspace{1cm} (15)  

\[ F(t) = \left( 1 - e^{-bt^\alpha} \right) \] \hspace{1cm} (16)  

Note that \( b > 0, t \in [0, \infty] \).

In the Weibull distribution such as Eq. 15 and Eq. 16, an exponential distribution is obtained when the shape parameter(\( \alpha \)) is 1, and a Rayleigh distribution is obtained when the shape parameter(\( \alpha \)) is 2. Therefore, the mean value function and the intensity function of the finite fault NHPP Rayleigh model are as follows.

\[ m(t, \theta, b) = \theta F(t) = \theta(1 - e^{-bt^2}) \] \hspace{1cm} (17)  

\[ \lambda(t, \theta, b) = \theta f(t) = 2\theta^2te^{-bt^2} \] \hspace{1cm} (18)  

Note that \( \theta > 0, b > 0 \).
The log-likelihood function to Maximum Likelihood Estimation (MLE) by using the Eq. 17 and Eq. 18 is derived as follows.

$$\ln L_{NHPP}(\theta | x) = n\ln 2 + n\ln \theta + n\ln b + \sum_{i=1}^{n} \ln x_i$$  \hspace{1cm} (19)

$$-b \sum_{i=1}^{n} x_i^2 - \theta (1 - e^{-bx_i^2})$$

Note that $x = 0 \leq x_1 \leq x_2 \leq \ldots \leq x_n$, $\theta$ is parameter space.

The partial derivatives of the parameters $\theta$ and $b$ are as follows.

$$\frac{\partial \ln L_{NHPP}(\theta | x)}{\partial \theta} = \frac{n}{\theta} - 1 + e^{-bx_i^2} = 0 \hspace{1cm} (20)$$

$$\frac{\partial \ln L_{NHPP}(\theta | x)}{\partial b} = \frac{n}{b} - \sum_{i=1}^{n} x_i^2 - \theta x_i^2 e^{-bx_i^2} = 0 \hspace{1cm} (21)$$

Note that $x = (x_1, x_2, x_3 \ldots x_n)$.

Therefore, the maximum likelihood estimator $\hat{\theta}_{MLE}$ and $\hat{b}_{MLE}$ satisfying the following the Eq. 20 and Eq. 21 can be estimated by a numerical method.

### 2.4 Finite Failure NHPP: Inverse-Exponential Model

The Inverse-weibull distribution is a widely applied distribution in reliability and medical field. In this Inverse-weibull distribution, the Inverse-exponential distribution is obtained when the shape parameter ($\gamma$) is 1. Here, the cumulative distribution function $F(t)$ of the Inverse-weibull distribution is as follows[10].

$$F(t) = e^{-(bt)^{-\gamma}} \hspace{1cm} (22)$$

Note that $b > 0$, $t \in [0, \infty]$, $\gamma$ is shape parameter.

Therefore, when the shape parameter conditions ($\gamma = 1$) are applied in the Inverse-weibull distribution, the probability density function $f(t)$ and the cumulative distribution function $F(t)$ of the Inverse-exponential distribution are as follows.

$$f(t) = F(t)' = b^{-1}t^{-2}e^{-(bt)^{-1}} \hspace{1cm} (23)$$

$$F(t) = e^{-(bt)^{-1}} \hspace{1cm} (24)$$

Note that $b > 0$, $t \in [0, \infty]$.

Therefore, the mean value function and the intensity function of the finite fault NHPP Inverse-exponential model are as follows.

$$m(t | \theta, b) = \theta F(t) = \theta e^{-(bt)^{-1}} \hspace{1cm} (25)$$

$$\lambda(t | \theta, b) = \theta f(t) = \theta b^{-1}t^{-2}e^{-(bt)^{-1}} \hspace{1cm} (26)$$

The log-likelihood function to Maximum Likelihood Estimation (MLE) by using the Eq. 25 and Eq. 26 is derived as follows.

$$\ln L_{NHPP}(\theta | x) = n\ln \theta - n\ln b$$

$$+ 2 \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} (bx_i) - 1 - \theta e^{-(bx_i)^{-1}} = 0 \hspace{1cm} (27)$$

The partial derivatives of the parameters $\theta$ and $b$ are as follows.

$$\frac{\partial \ln L_{NHPP}(\theta | x)}{\partial \theta} = -\frac{n}{\theta} - e^{-(bx_i)^{-1}} = 0 \hspace{1cm} (28)$$

$$\frac{\partial \ln L_{NHPP}(\theta | x)}{\partial b} = \frac{n}{b} + 1 - \frac{bx_i}{b^2} e^{-(bx_i)^{-1}} = 0 \hspace{1cm} (29)$$

Note that $x = (x_1, x_2, x_3 \ldots x_n)$.

Therefore, the maximum likelihood estimator $\hat{\theta}_{MLE}$ and $\hat{b}_{MLE}$ satisfying the following the Eq. 28 and Eq. 29 can be estimated by a numerical method.

### 3. The Proposed Analysis Algorithm and Solutions

The analysis algorithm of the proposed software reliability model is as follows.

Step 1: Validating the software failure data collected through the Laplace trend test analysis.

Step 2: Calculating the parameters ($\hat{\theta}, \hat{b}$) for the proposed model using the Maximum Likelihood Estimation.

Step 3: Calculating coefficient of determination ($R^2$) and Mean Square Error (MSE) for efficient model selection.

Step 4: Analyzing the attributes ($m(t), \lambda(t)$) and future reliability($R(t)$) of proposed models.

Step 5: Providing research information on the optimal model by analyzing performance of the proposed model.

After analyzing the performance of the proposed model using the above steps, we will present information on the model that software developers need.

Let compare and analyze the performance of the proposed reliability models using the software failure time data[11] as shown in Table 1. This software failure is the data that was occurred 30 times in 187.35 unit time.
Laplace trend test was used to verify the reliability of the software failure time data as shown in Fig 1.

![Laplace trend test](image)

**Table 1.** Software Failure Time Data

<table>
<thead>
<tr>
<th>Failure number</th>
<th>Failure time (hours)</th>
<th>Failure interval-time</th>
<th>Failure time (hours)×10⁻¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.79</td>
<td>4.79</td>
<td>0.479</td>
</tr>
<tr>
<td>2</td>
<td>7.45</td>
<td>2.66</td>
<td>0.745</td>
</tr>
<tr>
<td>3</td>
<td>10.22</td>
<td>2.77</td>
<td>1.022</td>
</tr>
<tr>
<td>4</td>
<td>15.76</td>
<td>5.54</td>
<td>1.576</td>
</tr>
<tr>
<td>5</td>
<td>26.10</td>
<td>10.34</td>
<td>2.610</td>
</tr>
<tr>
<td>6</td>
<td>35.59</td>
<td>9.49</td>
<td>3.559</td>
</tr>
<tr>
<td>7</td>
<td>42.52</td>
<td>6.93</td>
<td>4.252</td>
</tr>
<tr>
<td>8</td>
<td>48.49</td>
<td>5.97</td>
<td>4.849</td>
</tr>
<tr>
<td>9</td>
<td>49.66</td>
<td>1.17</td>
<td>4.966</td>
</tr>
<tr>
<td>10</td>
<td>51.36</td>
<td>1.70</td>
<td>5.136</td>
</tr>
<tr>
<td>11</td>
<td>52.53</td>
<td>1.17</td>
<td>5.253</td>
</tr>
<tr>
<td>12</td>
<td>65.27</td>
<td>12.74</td>
<td>6.527</td>
</tr>
<tr>
<td>13</td>
<td>69.96</td>
<td>4.69</td>
<td>6.996</td>
</tr>
<tr>
<td>14</td>
<td>81.70</td>
<td>11.74</td>
<td>8.170</td>
</tr>
<tr>
<td>15</td>
<td>88.63</td>
<td>6.93</td>
<td>8.863</td>
</tr>
<tr>
<td>16</td>
<td>107.71</td>
<td>19.08</td>
<td>10.771</td>
</tr>
<tr>
<td>17</td>
<td>109.06</td>
<td>1.35</td>
<td>10.906</td>
</tr>
<tr>
<td>18</td>
<td>111.83</td>
<td>2.77</td>
<td>11.183</td>
</tr>
<tr>
<td>19</td>
<td>117.79</td>
<td>5.96</td>
<td>11.779</td>
</tr>
<tr>
<td>20</td>
<td>125.36</td>
<td>7.57</td>
<td>12.536</td>
</tr>
<tr>
<td>21</td>
<td>129.73</td>
<td>4.37</td>
<td>12.973</td>
</tr>
<tr>
<td>22</td>
<td>152.03</td>
<td>22.30</td>
<td>15.203</td>
</tr>
<tr>
<td>23</td>
<td>156.40</td>
<td>4.37</td>
<td>15.640</td>
</tr>
<tr>
<td>24</td>
<td>159.80</td>
<td>3.40</td>
<td>15.980</td>
</tr>
<tr>
<td>25</td>
<td>163.85</td>
<td>4.05</td>
<td>16.385</td>
</tr>
<tr>
<td>26</td>
<td>169.60</td>
<td>5.75</td>
<td>16.960</td>
</tr>
<tr>
<td>27</td>
<td>172.37</td>
<td>2.77</td>
<td>17.237</td>
</tr>
<tr>
<td>28</td>
<td>176.00</td>
<td>3.63</td>
<td>17.600</td>
</tr>
<tr>
<td>29</td>
<td>181.22</td>
<td>5.22</td>
<td>18.122</td>
</tr>
<tr>
<td>30</td>
<td>187.35</td>
<td>6.13</td>
<td>18.735</td>
</tr>
</tbody>
</table>

In general, if the Laplace factor estimates are distributed between -2 and 2, the data are reliable because the extreme values do not exist and are stable.

As a result of this test in this Fig 1, the estimated value of the Laplace factor was distributed between 0 and 2, as shown in Fig. 1. Therefore, it is possible to apply this data because there is no extreme value[12].

In this study, the Maximum Likelihood Estimation (MLE) was used to perform parameter estimation. And numerical conversion data (Failure time[hours] × 10⁻¹) in order to facilitate the parameter estimation was used. The calculation method of the nonlinear equations is solved using the bisection method, and the results are shown in Table 2.

**Table 2.** Parameter Estimation of Each Model

<table>
<thead>
<tr>
<th>Model</th>
<th>MLE</th>
<th>Model comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSE</td>
<td>R²</td>
</tr>
<tr>
<td>Goel-Okumoto</td>
<td>( \hat{\theta} = 32.9261 ) ( \hat{b} = 0.1297 )</td>
<td>32.9379 0.8956</td>
</tr>
<tr>
<td>Rayleigh</td>
<td>( \hat{\theta} = 30.0412 ) ( \hat{b} = 0.0188 )</td>
<td>32.1798 0.8980</td>
</tr>
<tr>
<td>Inverse-exponential</td>
<td>( \hat{\theta} = 41.2881 ) ( \hat{b} = 0.1692 )</td>
<td>20.2035 0.9359</td>
</tr>
</tbody>
</table>

Explanatory notes.

MLE = Maximum Likelihood Estimation

MSE = Mean Square Error, \( R^2 = Coefficient of Determination \).
As the basis for determining the efficient model, the mean square error (MSE) is defined as follows:

\[
MSE = \frac{1}{n-k} \sum_{i=1}^{n} (m(x_i) - \hat{m}(x_i))^2
\]  

(30)

Note that \(m(x_i)\) is the total accumulated number of errors observed within time \((0, x_i)\), \(\hat{m}(x_i)\) is the estimated cumulative number of errors at time \(x_i\) obtained from the fitting mean value function, \(n\) is the number of observations and \(k\) is the number of parameters to be estimated. In efficient model selection, the smaller the mean square error, the more efficient the model.

The coefficient of determination \((R^2)\) is a measuring value to the explanatory power of the difference between the target value and the observed value. The larger the value of the decision coefficient in efficient model selection, the more efficient the model because the error is relatively small. It is defined as

\[
R^2 = 1 - \frac{\sum_{i=1}^{n} (m(x_i) - \hat{m}(x_i))^2}{\sum_{i=1}^{n} (m(x_i) - \bar{m}(x_i))^2}
\]  

(31)

As shown in Table 2, we can see that the Rayleigh model is more efficient than the Goel-Okumoto model. But, the Inverse-exponential model has the largest coefficient of determination and the smallest mean square error is more efficient than the other models[13].

Also, Fig 2 shows the transition of mean square error (MSE) according to each failure number. That is, in this figure, the Inverse-exponential model shows better estimates than the other model in the total range of failure number.

Let us analyze the reliability performance of the proposed models for future mission time. Here, reliability is the probability that a software failure will occur when testing at \(x_n = 18.735\), and no software failure will occur between confidence intervals \([x_n, x_n + \tau]\) where \(\tau\) is the future mission time.

Therefore, the reliability of future mission time is as follows[14].
\[
R(t|x_n) = e^{-\int_{x_n}^{t} \lambda(\tau)d\tau} = \exp[-(m(t + x_n) - m(x_n))]
\]

As shown in Fig 5, the Inverse-exponential model and the Rayleigh model show a higher reliability trend than the Goel-Okimoto model in which the reliability decreases with the mission time.

4. CONCLUSION

It is possible to efficiently improve the reliability performance by analyzing the performance after quantitatively modeling the occurrence of the failure in the software test operation or the software development process. In this study, based on the finite-fault NHPP model with software failure time data, we compared and analyzed the software reliability performance of the Inverse-exponential model and the Rayleigh model which is Weibull life distribution, together with Goel-Okimoto basic model.

The results of this study can be summarized as follows.

First, in the performance analysis of the strength function, Inverse-exponential model was effective because the inverse-exponential model decreased along with the Rayleigh model as the failure time passed, and the mean square error (MSE) showed the smallest trend.

Second, in the performance analysis of the mean value function, all the proposed models showed overestimation patterns in the error estimation for true values, but the Inverse-exponential model was the most efficient because it had the smallest error value than the other models.

Third, in the performance analysis of mission reliability, the Inverse-exponential model and the Rayleigh model show stable and high reliability trends. On the other hand, Goel-Okimoto basic model shows that reliability decreases with the failure time. In other words, a comprehensive analysis of these simulations results shows that the inverse-exponential model is the most efficient model with the best performance among the proposed models.

As a result, through this study, along with a new analysis on the reliability performance of the proposed model without existing research examples, we were able to provide the research information that software developers can use as basic design guideline.

In addition to, future research will be needed to find the optimal model through the reliability performance analysis after applying the same type of software failure time data to various reliability models.

ACKNOWLEDGEMENT

Funding for this paper was provided by Namseoul University.

REFERENCES


