A Performance Analysis of Software Reliability Model using Lomax and Gompertz Distribution Property

Hee-Cheul Kim*

Department of Industrial and Management Engineering, Namseoul University, S. Korea; kim1458@nsu.ac.kr

Abstract

Background/Objectives: In this research, the comparative problems of a reliability model about Lomax and Gomperz distribution property were proposed. The maximum likelihood estimation and bisection method were used to the parameter estimation. Methods/Statistical Analysis: As special occasions, the model selection based on the Mean Square Error and coefficient of determination for the efficient model was conducted. Analysis of a failure time using real data set from the proposing reliability model (based on Lomax and Gomperz distribution) was worked. To obtain for the data reliability, the Laplace trend test was employed. Findings: In this study, the proposed Lomax than Gomperz distribution model is more efficient model in this area. The case of the higher shaping parameter of the Lomax distribution model is judged more reliable model in this field. Thus, Lomax distribution model can also be applied as a special model. Improvements: The software testing for the debugging to reduce cost in terms of the reliability from software is essential problem. From a research, the software developers must be considered for the growth model by the prior knowledge of the software to identify failure modes which can be able to help.

Keywords: Gomperz Distribution, Laplace Trend Test, Lomax Distribution, Mission Time, NHPP

1. Introduction

Software reliability is the probability that can operate without failure for a period of time at constant born environmental conditions. Thus, the reliability from the software is a central difficult during the development process of the software. The subject must be satisfied the requirements of the user and testing cost. The reliability and the cost during the development process of software with considerations for a release time can be an indispensable problem.

Many software reliability models have been accomplished so far. These models of the software reliability, using Non-Homogenous Poisson Process (NHPP), have the assumption that if a fault occurs, immediately removed and no new fault has occurred during the debugging process.

In this software reliability field, the Enhanced Non-Homogenous Poisson Process (ENHPP) model was presented using test coverage property. In proposed software reliability models with the property of an exponential curve. The feature type of the mean value function was applied to the total number of defects have S-shaped or exponential-shaped form. The generalized models using a growth model of the delayed S-shaped reliability and inflection S-shaped reliability were presented. In proposed a software reliability changing point problem and generalized reliability growth model using changing point was presented. An evaluation of software stability was presented by. In conjunction with this model presented an efficient technique to predict a software reliability using the testing-effort function of generalized logistic property and the parameter for change-point form. In described that S-type model can be applied to the learning efficient for the software failure number and test tools.

In this study, the characteristics of Lomax and Gompertz distribution NHPP model for the software reliability were employed using a nonhomogeneous Poisson process from the faults of the infinite number.

* Author for correspondence
2. Assist Work

2.1 NHPP Model

The Non-Homogeneous Poisson Process (NHPP) model consists of a mean value function and intensity function. These factors must be satisfied the following properties:

\[ m(t) = \int_0^t \lambda(t)\, dt \quad \frac{dm(t)}{dt} = \lambda(t) \quad (1) \]

In Equation (1), \( N(t) \) follows Probability Density Function (PDF) of Poisson distribution using the element \( m(t) \). Applying these conditions, the Poisson probability density function can be represented next pattern.

\[ p(N(t) = n) = \frac{[m(t)]^n}{n!} e^{-m(t)} \quad n = 0, 1, \ldots, \infty \quad (2) \]

These time-area models based on NHPP were distinguished using a probability of the failure-time. In Equation (2), the mean value function \( m(t) \) is parameter. The failure intensity function \( \lambda(t) \) can be characterized using the relationship of Equation (1).

A finite model has the assumption no new fault has occurred at the time of each repair. But, new failure at the time of repair from the real circumstances may happen.

To reflect this state of affairs, a NHPP model using the RVS (Record Value Statistics) can be used. The mean value function was known to next condition.

\[ m(t) = -\ln(1 - F(t)) \quad (3) \]

A pattern of the Equation (3) is mean value function about the infinite failure NHPP model.

The form for the intensity function using the relationship of Equation (1) can be derived the hazard function \( h(t) \). Namely,

\[ \lambda(t) = m'(t) = f(t)/(1 - F(t)) = h(t) \quad (4) \]

From Equation (4), \( f(t) \) denotes PDF (Probability Density Function) and \( F(t) \) is CDF (Cumulative Distribution Function).

Let \( \{t_n, n = 1, 2, \ldots\} \) indicate order statistic of the times among sequential software failure. So, \( t_n \) represents the difference for failure time from \((n-1)^{th}\) failure to \(n^{th}\) failure. Consequently, the failure last time \( x_n \) represents \( n^{th}\) failure time and can be expressed by the next relationship.

\[ x_n = \sum_{i=1}^{n} t_i \quad (i = 1, 2, \ldots, n; \quad 0 < x_1 < x_2 < \cdots < x_n) \quad (5) \]

The likelihood function (in some cases, joint Probability Density Function) of \( x_1, x_2, \ldots, x_n \) was given to next form:

\[ f(x_1, x_2, \ldots, x_n) = L(\theta) = e^{-m(x_n)} \prod_{i=1}^{n} \lambda(x_i) \quad (6) \]

In Equation (6), \( x=(x_1, x_2, x_3, \ldots, x_n) \) and \( \Theta \) denotes the parameter space.

For realizations of the random variables \((x_1, x_2, \ldots, x_n)\) for given order statistic of the software failure times \( (x_i, x_{i+1}, \ldots, x_n) \), the parameter estimation of the NHPP property about the software reliability can be calculated using the method of Maximum Likelihood Estimation (MLE).

In these circumstances, the reliability for conditional form \( \hat{R} (\xi | x_n) \) can be expressed \( R(t|s) \) next equation form.

\[ \hat{R}(\xi | x_n) = e^{-\int_{x_n}^{\xi} \lambda(t)\, dt} = \exp[-(m(x_n) + \xi - m(x_n))] \quad (7) \]

In Equation (7), \( \xi \) means the work time and \( x_n \) denotes the latest failure-time.

2.2 Lomax Distribution

The Lomax distribution provisionally also termed the Pareto Type II distribution, is a weighty-tail probability distribution often applied in commerce, finances and actual affair. It is called after K. S. Lomax. It is fundamentally a Pareto distribution that has been lifted so that its provision activates at zero.

The Probability Density Function (PDF) and distribution function for the Lomax distribution were known to next forms.

\[ f(t) = \frac{\lambda k}{(1 + \lambda t)^{k+1}}, \quad t > 0 \quad (8) \]

\[ F(t) = 1 - (1 + \lambda t)^{-k} \quad (9) \]

Note that \( k > 0 \) is shape parameter and \( \lambda > 0 \) is scale parameter. The hazard function \( h(t) \) can be derived as follows.

\[ h(t) = \frac{f(t)}{1 - F(t)} = \frac{\lambda k}{(1 + \lambda t)} \quad (10) \]

Some of the typical Lomax density functions for the different value of \( k > 0 \) and for \( \lambda = 0.5 \) are provided in Figure 1.
2.3 Comperz Distribution

The Gomperz distribution is a probability distribution from the continuous property and is widely used to describe the distribution of adult deaths. Also, Gomperz distribution has been applied in various fields of an industry distribution. The probability density function and distribution function for Gomperz distribution are known as follows:

\[ f(t | \alpha, \beta) = \alpha \beta e^{\beta t} e^{-\alpha e^{\beta t}} \]  \hspace{1cm} (11)

\[ F(t | \alpha, \beta) = 1 - \exp(-\alpha (e^{\beta t} - 1)) \]  \hspace{1cm} (12)

In Equation (11) and (12), \( \alpha > 0 \) denotes the shape parameter and \( \beta > 0 \) means the scale parameter and time space is \( t \in (0, \infty) \).

The hazard function and intensity are same form for an infinite failure-NHPP model. This hazard function can be expressed next form using Equation (11) and Equation (12).

\[ h(t) = \frac{F'(t)}{1 - F(t)} = \frac{f(t)}{1 - F(t)} = \alpha \beta e^{\beta t} = \lambda(t) \]  \hspace{1cm} (13)

2.4 Tools of the Model Comparison using Real Data Set

The judgment principles for the efficiency of the suggested model used to the Mean Square Error (MSE) and R square (\( R^2 \)).

In terms of the deviation for among forecast estimations and the real values, a Mean Square Error used and was known next form.

\[ MSE = \frac{\sum_{i=1}^{n} (m(x_i) - \hat{m}(x_i))^2}{n - k} \]  \hspace{1cm} (14)

In Equation (14), \( m_i \) means the whole cumulated number of the practical failure among time \((0, x_i]\) and \( \hat{m}_i \) is the estimated cumulative number of the failure at time \( x_i \). Also, \( n \) means the number of samples and \( k \) denotes the number of parameters to be estimated.

The \( R^2 \) can be used another tool. This tool can be estimated the quantity how effective fit from the variation of estimated values. This tool was known to next form.

\[ R^2 = 1 - \frac{\sum_{i=1}^{n} (m(x_i) - \hat{m}(x_i))^2}{\sum_{i=1}^{n} (m(x_i) - \bar{m})^2} \]  \hspace{1cm} (15)

2.5 Software Reliability Infinite NHPP Model based on Lomax and Gomperz Distribution

The mean value function \((m(t))\) and failure intensity function \((\lambda(t))\) of Lomax distribution model grounded on Equation (3) and Equation (4) were stated next forms.

\[ \lambda(t) = h(t) = \frac{f(t)}{1 - F(t)} = \frac{\lambda k}{(1 + \lambda t)} \]  \hspace{1cm} (16)

\[ m(t) = -\ln(1 - F(t)) = k \ln(1 + \lambda t) \]  \hspace{1cm} (17)

The log likelihood function from Lomax model based on Equation (6) can be derived as next equation.

\[ \ln L_{NHPP}(\hat{\lambda}, k | x) = n \ln \lambda + n \ln k - \sum_{i=1}^{n} \ln(1 + \lambda x_i) - k \ln(1 + \lambda x_n) \]  \hspace{1cm} (18)

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

For maximizing Equation (18), the fixed values of \( k(30, 40, 50) \) with respect to \( \lambda \) must be satisfied next condition:

\[ \frac{\partial \ln L_{NHPP}(\lambda, k | x)}{\partial \lambda} = -\frac{n}{\lambda} - \sum_{i=1}^{n} \frac{x_i}{1 + \lambda x_i} - k \left( \frac{x_n}{1 + \lambda x_n} \right) = 0 \]  \hspace{1cm} (19)

Using \( m(x_n + \delta) = k \ln(1 + \lambda (\delta + x_n)) \) and \( m(x_n) = k \ln(1 + \lambda x_n) \), the reliability can be estimated using the next form.
\hat{R}(\delta \mid x_n) = \exp[-\{m(\delta + x_n) - m(x_n)\}] \tag{20}

Note that \(m(x_n + \delta)\) and \(m(x_n)\) are the mean value time for the mission time \(\delta\) and \(x_n\) is the failure-last time.

Also, the mean value function \(m(t)\) and failure intensity function \(\lambda(t)\) for Comperz software reliability model created on Equation (11) and Equation (12) can be obtained next equations.

\[
\lambda(t) = \frac{f(t)}{1 - F(t)} = \alpha \beta e^{\beta t} = h(t) \tag{21}
\]

\[
m(t) = -\ln(1 - F(t)) = \alpha (e^{\theta t} - 1) \tag{22}\]

Similarly, the likelihood function for Gompertz model must be satisfied next condition.

\[
L_{NHPP}(\alpha, \beta \mid x) = \left(\prod_{i=1}^{n} \alpha \beta e^{\beta x_i}\right) \exp[-\alpha (e^{\beta x_i} - 1)] \tag{23}\]

The maximum likelihood estimation for the parameter estimation was used. Therefore, the log likelihood function based on Equation (23) is can be expressed as follows.

\[
\ln L_{NHPP}(\Theta \mid x) = n \ln \alpha + n \ln \beta - \beta \sum_{i=1}^{n} \ln x_i - \alpha (e^{\beta x_i} - 1) \tag{24}\]

In Equation (24), \(\hat{\alpha}_{MLE}\) and \(\hat{\beta}_{MLE}\) must be satisfied the following conditions for the parameter estimation.

\[
\frac{\partial \ln L_{NHPP}(\Theta \mid x)}{\partial \alpha} = \frac{n}{\alpha} - (e^{\beta x_i} - 1) = 0 \tag{25}\]

\[
\frac{\partial \ln L_{NHPP}(\Theta \mid x)}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^{n} x_i - \alpha x_i e^{\beta x_i} = 0 \tag{26}\]

Similarly, using \(m(\delta + x_n) = \alpha (e^{\beta (\delta + x_n)} - 1)\) and \(m(x_n) = \alpha (e^{\beta x_n} - 1)\), the reliability can be estimated as follows:

\[
\hat{R}(\delta \mid x_n) = \exp[-\{m(\delta + x_n) - m(x_n)\}] \tag{27}\]

Note that \(m(x_n + \delta)\) and \(m(x_n)\) are the mean value time for the mission time \(\delta\) and \(x_n\) denotes the failure last time.

### 3. Illustration of Failure Time

In this chapter, a software failures time data\(^{13}\) used to analyze the features of shaping parameter from the Lomax and Gompertz distribution model. This data set lists in Table 1.

<table>
<thead>
<tr>
<th>Fault Number</th>
<th>Time to failure (hours)</th>
<th>Fault number</th>
<th>Time to failure (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.479</td>
<td>16</td>
<td>10.771</td>
</tr>
<tr>
<td>2</td>
<td>0.745</td>
<td>17</td>
<td>10.906</td>
</tr>
<tr>
<td>3</td>
<td>1.022</td>
<td>18</td>
<td>11.183</td>
</tr>
<tr>
<td>4</td>
<td>1.576</td>
<td>19</td>
<td>11.779</td>
</tr>
<tr>
<td>5</td>
<td>2.61</td>
<td>20</td>
<td>12.536</td>
</tr>
<tr>
<td>6</td>
<td>3.559</td>
<td>21</td>
<td>12.973</td>
</tr>
<tr>
<td>7</td>
<td>4.252</td>
<td>22</td>
<td>15.203</td>
</tr>
<tr>
<td>8</td>
<td>4.849</td>
<td>23</td>
<td>15.64</td>
</tr>
<tr>
<td>9</td>
<td>4.966</td>
<td>24</td>
<td>15.98</td>
</tr>
<tr>
<td>10</td>
<td>5.136</td>
<td>25</td>
<td>16.385</td>
</tr>
<tr>
<td>11</td>
<td>5.253</td>
<td>26</td>
<td>16.96</td>
</tr>
<tr>
<td>12</td>
<td>6.527</td>
<td>27</td>
<td>17.237</td>
</tr>
<tr>
<td>13</td>
<td>6.996</td>
<td>28</td>
<td>17.6</td>
</tr>
<tr>
<td>14</td>
<td>8.17</td>
<td>29</td>
<td>18.122</td>
</tr>
<tr>
<td>15</td>
<td>8.863</td>
<td>30</td>
<td>18.735</td>
</tr>
</tbody>
</table>

To determine the validity of data, the trend test\(^{9}\) should be preceded. In this paper, the Laplace trend test analysis is used for the validity of the data. The result of the Laplace trend test is given to Figure 2. In this figure, because the value of the Laplace factor is indicated between 2 and -2, it is adequate to estimate the reliability\(^{14}\) using this data.

![Figure 2. Laplace trend test.](image)
accomplished iteration of 50 times.

A result of the parameter estimation was listed in Table 2. In this table, the Mean Square Error (MSE), coefficient of determination ($R^2$) and maximum likelihood estimation were listed for the and Lomax model.

In Table 2, for the software model contrast, MSE (which measures the modification for among the actual value and forecasted value) shows that a case of the shaping parameter $k=50$ (from Lomax model) than $k=40$ (from Lomax model) and Gompertz model has smaller value. Therefore, a case of the shaping parameter $k=50$ is appreciably better than other shaping parameters (from Lomax model) and the Gompertz model. Similarly, $R^2$ (which means that the predictive power of the difference for the between predicted values) show that a case of the shaping parameter $k=50$ (from Lomax model) model than other cases has higher value. Thus, the case of the shaping parameter $k=50$ (from Lomax model) model than other shaping parameters is the utility model. Eventually, in terms of the deviation about among of the predicted values from the actual values, the shaping parameter $k=50$ (from Lomax model) model regard as the best model and in terms of the predictive power about the difference for the between predicted values, the shaping parameter $k=50$ (from Lomax model) model is the best model.

<table>
<thead>
<tr>
<th>Table 2. Laplace trend test.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Gompertz Model</td>
</tr>
<tr>
<td>Lomax Model</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Note. MLE: Maximum Likelihood Estimation; MSE: Mean Square Error; $R^2$: Coefficient of determination

The estimated results of mean value functions were listed in Figure 3. The estimated patterns of mean value function have the trends of the non-decreasing arrangement.

Also, the estimated results of the intensity function were listed in Figure 4. The estimated patterns of the intensity functions from Lomax model have the non-increasing arrangement and Gompertz model have the trends of the non-decreasing arrangement.

In Figure 5, in terms of comparison for the reliability, the case of the reliability for the assumed mission time show that the case of the shaping parameter $k=30$ (from Lomax model) than shaping parameter $k=40$ (from Lomax model), the shaping parameter $k=50$ (from Lomax model) and the Gompertz model was shown the highest reliability. Namely, the reliability has for the sensitive property to the mission time. As the reliability result, a case of the smaller shaping parameter from the Lomax distribution model is judged more reliable model in this field.
4. Conclusions

A special feature of the software reliability conformation with the Lomax and Gompertz distribution property that made out efficiency about application for software reliability was proposed. In terms of deviation for the among forecasted values and the actual observations, the higher shaping model of the Lomax model regard as the best model; while in terms of the predictive power of the difference for the between predicted values, the higher shaping parameter value of the Lomax distribution is the best model. The estimated result of the mean value functions has the tendency of the non-decreasing form. The estimated results of the intensity functions have the non-increasing form and the Gompertz model has the tendency of the slight increasing form.

The case of the higher shaping parameter of the Lomax distribution model is judged more reliable model in this field. This study can be given advance information to the software developer.

5. Acknowledgment

Funding for this paper was provided by Namseoul University.

6. References