International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 12, Number 6 (2019), pp. 760-764

A Comparative Study on the Finite Failure Software Reliability Model with Modified Lindley Type Lifetime Distribution

Hee-Cheul Kim[1]
[1]*Associate Professor,*
Department of Industrial & Management Engineering,
Namseoul University, South Korea.

Abstract

In this paper, were analysed the software failure time data using the lifetime distribution following existing Goel-Okumoto model and the modified Lindley distribution. The software reliability model based on the non-homogeneous Poisson process with the finite number of failures was discussed. Therefore, this study was used to compare and analysed the software failure characteristics using mean squared error, mean value function and hazard function as a measure to compare the reliability characteristics. The results of this study show that the Goel-Okumoto model is smaller than the modified Lindley distribution model and the mean value function pattern is similar to the true value, but the Goel-Okumoto model is smaller to the modified Lindley distribution. Therefore, the existing Goel-Okumoto model can be regarded as an efficient model. However, in the form of the reliability function have the non-incremental pattern as mission time passes and the model with the modified Lindley distribution are higher than those of the Goel-Okumoto model. Through this study, software operators can use the mean square error, average value, and hazard function trends to identify the types of failures in software reliability that reflect the characteristics of various lifetime.

Keywords: Software Reliability Model, Modified Lindley distribution, NHPP, Hazard function, NHPP

I. INTRODUCTION

Software operating systems have long been an indispensable tools of our lives. The reliability of such an operating system can be a fundamental factor in the software operating environment because it can provide better quality service to software users. However, the development of a software operating system can be regarded as a complicated and difficult process due to time and cost. Therefore, software operations and developers aim to improve the stability of software systems. Software reliability engineering research has been proposed for a long time. These models have been proposed a software reliability model based on the non-homogeneous Poisson process (NHPP) using the failure intensity function and the mean value function to determine the reliability characteristics such as number of failures and failure rate [1]. Thus, the software reliability model is used to analyse the software failure phenomenon by estimating the reliability of the software, the remaining number of failures, the failure intensity and the software development cost. Yamada and Osaki [2] were studied various software stability features by estimating the results of the mean value function using the maximum likelihood method and Teng and Pham [3] were used a generalized software reliability model to determine the characteristics that affect the software failure rate. Hee-Cheul Kim [4] also was studied reliability characteristics using Burr-XII and Type-2 Gumbel life time distributions. This study was compared the reliability characteristics of the software reliability model based on the non-homogeneous Poisson process with finite number of failures using Goel-Okumoto model and the modified Lindley distribution developed to analysed the failure time data-

II. FINITE NHPP SOFTWARE RELIABILITY MODEL

2.1. Existing Goel-Okumoto

The most basic model in this field is the Goel-Okumoto model. This model assumes the exponential distribution as the life time distribution per fault. Therefore, the rate of occurrence of faults is constant and the intensity function and the average value function are known as follows [5].

$$\lambda(t\,|\,\theta,\beta)=\theta f(t)=\theta\beta e^{-\beta t} \tag{1}$$

$$m(t\,|\,\theta,\beta)=\theta F(t)=\theta(1-\beta e^{-\beta t}) \tag{2}$$

In finite failure NHPP model, $t\in(0,\infty]$ and θ was specified the expected value of faults that would be discovered observing time $(0,t]$. The shape parameter β is failure rate, $f(t)$ is probability density function and $F(t)$ is cumulative distribution. In equation (2), time t and x_n are replaced with the last failure time point, the likelihood function is known as follows.

$$\begin{aligned} L_{NHPP}(\Theta|\underline{x}) &= \left[\prod\nolimits_{i=1}^{n} \lambda(x_i)\right]\times\exp\left[-m(x_n)\right] \\ &= \left[\prod\nolimits_{i=1}^{n} \theta\beta e^{-\beta x_i}\right]\times\exp\left[-\theta(1-\beta e^{-\beta x_n})\right] \end{aligned} \tag{3}$$

Note. $\underline{x}=(x_1\le x_2\le x_3\le\ldots\le x_n)$,
$\Theta=\{\theta,\beta\}$ Specifies parameter space.

The log-likelihood function by means of the equation (3) can be detailed ensuing relation [3, 6].

$$\ln L_{NHPP}(\Theta|\underline{x}) = -m(x_n) + \sum_{i=1}^{n} \ln \lambda(x_i) \\ = n\ln\theta + n\ln\beta - \beta\sum_{i=1}^{n} x_i - \theta(1-\beta e^{-\beta x_n}) \tag{4}$$

The estimator $\hat{\theta}_{MLE}$ and $\hat{\beta}_{MLE}$ must be assessed the following structure for the maximum likelihood estimation about all parameter by means of the equation (4).

$$\frac{\partial \ln L_{NHPP}(\Theta|\underline{x})}{\partial\theta} = \frac{n}{\theta} - 1 + e^{-\beta x_n} = 0 \tag{5}$$

$$\frac{\partial \ln L_{NHPP}(\Theta|\underline{x})}{\partial\beta} = \frac{n}{\beta} - \sum_{i=1}^{n} x_n - \theta x_n e^{-b x_n} = 0 \tag{6}$$

The estimator $\hat{\theta}_{MLE}$ and $\hat{\beta}_{MLE}$ using equation (5) and (6) can be summarized as follows.

$$\frac{n}{\hat{\theta}_{MLE}} = 1 - e^{-\beta x_n} \tag{7}$$

$$\frac{n}{\hat{\beta}_{MLE}} = \sum_{i=1}^{n} x_n + \hat{\theta}_{MLE}\, x_n\, e^{-\hat{\beta}_{MLE} x_n} = 0 \tag{8}$$

2.2. Model of Modified Lindley Type Lifetime Distribution

The Lindley distribution was introduced by Lindley [7] to analyze failure time data, especially in applications model for stress-strength reliability. The hazard function of the Lindley distribution is widely used because of its ability to model failure time data that follows increasing, decreasing, and bath-type patterns. The probability distribution function ($f(t)$) and the distribution function ($F(t)$) of the basic type (LM 1) of the Lindley distribution are as follows [7, 8].

$$f(t|b) = \frac{b^2}{(b+1)}(1+t)\times e^{-bt} \tag{9}$$

$$F(t|b) = 1 - \left[1 + \frac{bt}{(b+1)}\right]\times e^{-bt} \tag{10}$$

In finite failure NHPP model, θ was specified the expected value of faults that would be discovered observing time $(0, t]$. Note that $t \in (0,\infty]$ and $b > 0$ are the shape parameter. In finite failure NHPP model, the intensity function and the mean value function are known as follows [1, 2].

$$\lambda(t|\theta,b) = \theta f(t) = \theta\, \frac{b^2}{(b+1)}(1+t)\times e^{-bt} \tag{11}$$

$$m(t|\theta,\beta) = \theta F(t) = \theta\left[1 - \left(1 + \frac{bt}{(b+1)}\right)\times e^{-bt}\right] \tag{12}$$

The log-likelihood function by means of the equation (11) and (12) can be detailed ensuing relation [3].

$$\ln L_{NHPP}(\Theta|\underline{x}) = -\theta\left[1 - \left(1 + \frac{bt}{(b+1)}\right)\times e^{-bt}\right] \\ + n\ln\theta + 2n\ln b - n\ln(b+1) + \sum_{i=1}^{n}(1+x_i) - b\sum_{i=1}^{n} x_i \tag{13}$$

The estimator $\hat{\theta}_{MLE}$ and $\hat{b}_{MLE}$ must be assessed the following structure for the maximum likelihood estimation about all parameter by means of the equation (13).

Thus, the estimator $\hat{\theta}_{MLE}$ and $\hat{\beta}_{MLE}$ can be summarized as follows.

$$\frac{\partial \ln L_{NHPP}(\Theta|\underline{x})}{\partial\theta} = \frac{n}{\theta} - \left[1 - \left(1 + \frac{bt}{(b+1)}\right)\times e^{-bt}\right] = 0 \tag{14}$$

$$\frac{\partial \ln L_{NHPP}(\Theta|\underline{x})}{\partial b} = \\ -\theta x_n e^{-b x_n}\, \frac{(b^2 + b^2 x_n + 2b + b x_n)}{(b+1)^2} + \frac{2n}{b} - \frac{n}{b+1} - \sum_{i=1}^{n} x_n = 0 \tag{15}$$

In addition, the model modified by Shanker, R. [8] was presented. The probability distribution function and the distribution function for this modified model (LM 2) are as follows.

$$f(t|b) = \frac{b^2}{(b^2+1)}(b+t)\times e^{-bt} \tag{16}$$

$$F(t|b) = 1 - \left(1 + \frac{bt}{(b^2+1)}\right)\times e^{-bt} \tag{17}$$

Note. $t \in (0,\infty]$, $b > 0$ is the shape parameter.

In finite failure NHPP model, the intensity function and the mean value function are known as follows [1, 2].

$$\lambda(t|\theta,b) = \theta f(t) = \theta\, \frac{b^2}{(b^2+1)}(b+t)\times e^{-bt} \tag{18}$$

$$m(t|\theta,\beta) = \theta F(t) = \theta\left[1 - \left(1 + \frac{bt}{(b^2+1)}\right)\times e^{-bt}\right] \tag{19}$$

Similarly, the estimator $\hat{\theta}_{MLE}$ and $\hat{b}_{MLE}$ must be assessed the following construction for the maximum likelihood estimation about all parameter by means of the equation (18) and (19).

Thus, the estimator $\hat{\theta}_{MLE}$ and $\hat{\beta}_{MLE}$ can be summarized as follows.

$$\frac{\partial \ln L_{NHPP}(\Theta|\underline{x})}{\partial\theta} = \frac{n}{\theta} - \left[1 - \left(1 + \frac{bt}{(b^2+1)}\right)\times e^{-bt}\right] = 0 \tag{20}$$

$$\frac{\partial \ln L_{NHPP}(\Theta \mid \underline{x})}{\partial b} = -\theta x_n e^{-b x_n} \frac{(b^4 + b^3 x_n + 3b^2 + b x_n)}{(b^2+1)^2} + \frac{2n}{b} - \frac{2nb}{b^2+1} + \sum_{i=1}^{n} \frac{1}{b+x_i} - \sum_{i=1}^{n} x_n = 0 \quad (21)$$

III. SOFTWARE FAILURE TIME RELIABILITY ANALYSIS USING MODIFIED LINDLEY TYPE LIFE TIME DISTRIBUTION

Table 1. Failure time data

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

In this section, the reliability structures of the software reliability model were studied using the software failure time data [9]. The failure time data is revealed in Table 1. Furthermore, a trend test should be headed in order to assure reliability of data. In this study, the trend analysis was used was the Box-plot test [10]. Therefore, in Figure 1, since there is no data information that is out of the range between the upper limit (=15.64+1.5 $\times$ (15.04-4.0849) = 31.073) and the lower limit (=4.849-1.5 $\times$ (15.04-4.0849) = -11.584), it can be seen that no abnormal value or extreme value occurs. Thus, can confirm that the work of identifying the attributes of the reliability model by applying this data is stable [11, 12]. The parameter approximation was used to the traditional maximum likelihood method. In the calculating method of nonlinear equations, the bisection technique was used which is a numerical method. A result of the parameter approximation was reached from the Table 2. In this section, results of parameter estimate were itemized in Table 2. These controls solve the root exactly, since the initial values were specified 0.0001 and 1.000 and the tolerance value for the measurement of interval (10^{-5}) were specified, with an accomplished replication of 100 times using R-language [13] checking satisfactory convergent.

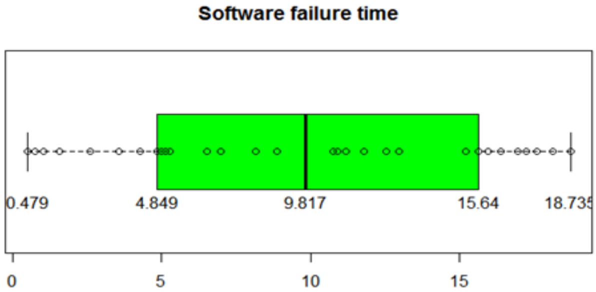

Fig. 1. Box plot test

Table 2. Parameter estimation of each model

Model	MLE	
Goel-Okumoto	$\hat{\theta}_{MLE} = 80.9562$	$\hat{\beta}_{MLE} = 2.417 \times 10^{-2}$
LM 1	$\hat{\theta}_{MLE} = 37.8877$	$\hat{b}_{MLE} = 1.497 \times 10^{-1}$
L M 2	$\hat{\theta}_{MLE} = 36.0896$	$\hat{b}_{MLE} = 1.703 \times 10^{-1}$
Model	Model Comparison	
	MSE	R^2
Goel-Okumoto	2.275	0.988
LM 1	4.618	0.979
L M 2	5.799	9.976

Note. LM 1 : the basic model of the Lindley distribution
LM 2 : the modified model of the Lindley distribution
MLE : Maximum likelihood estimation.
MSE : Mean square error.
R^2 : Coefficient of determination.

Based on the parameter estimates listed in Table 2, the estimated values of the mean square error (MSE) [11, 12], which is a measure of the difference between the actual value and the predicted value, are listed in Table 2 and the statistical tool of the mean square error [4 , 14] is as follows.

$$MSE = \frac{\sum_{i=1}^{n} [m(x_i) - \hat{m}(x_i)]^2}{n-k} \quad (22)$$

Note that $m(x_i)$ is the occupied cumulated number of the faults noticed in $(0, x_i]$ and $\hat{m}(x_i)$ approximating full cumulated number of the faults noticed in $(0, x_i]$, n specifies the number of realizing values and k is the number of the parameter. In Table 2, since the Goel-Okumoto model is smaller than the LM 1 and LM 2 models in the overall mean square error, the Goel-Okumoto model can be regarded as an efficient model for the LM 1 and LM 2 models and LM 1 is more efficient than the LM 2 model. In order to settle this situation, a summary picture of the comparison of estimated values of square error ($SE = [m(x_i) - \hat{m}(x_i)]^2, i = 1, 2 \ldots 30$) [4] for each failure time points are abridged in Figure 2. In this figure, squared error value of the Goel-Okumoto model shows a smaller than the LM 1 and LM 2 models as the failure time increases.

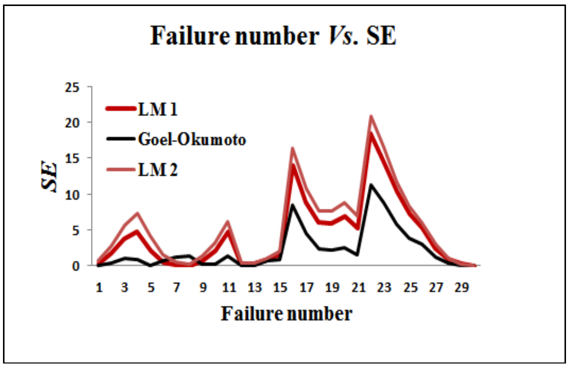

Fig. 2. Estimation of square error for each time

In addition, the coefficient of determination (R^2) is defined as follows as a tool for explaining the difference of predicted values. Therefore, a model with a larger coefficient of determination is considered an efficient model [10, 11].

$$R^2 = 1 - \frac{\sum_{i=1}^{n} [m(x_i) - \hat{m}(x_i)]^2}{\sum_{i=1}^{n} \left(m(x_i) - \sum_{j=1}^{n} m(x_j)/n \right)^2} \quad (23)$$

Thus, Goel-Okumoto model can be regarded as an efficient model because the estimated value for coefficient of determination in Table 2 has a larger estimated value than the LM 1 and LM 2 models. However, since the proposed model also has 95% or more, all models are considered to be efficient models [10, 11]. Figure 3 shows that the Goel-Okumoto model and the LM 1 and LM 2 models are almost similar to each other in terms of the mean value function pattern, but the Goel-Okumoto model has a relatively smaller width in terms of real value than the LM 1 and LM 2 models.

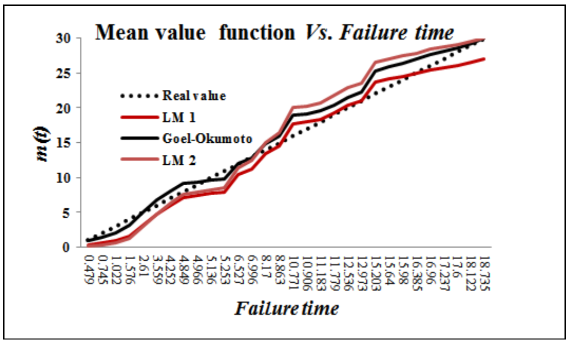

Fig. 3. Pattern of mean value function for each model

The hazard function, which means the instantaneous failure rate for the specified failure time, is defined by the following pattern [4].

$$h(t) = \frac{f(t)}{1 - F(t)} \quad (24)$$

Note that probability distribution function is $f(t)$ and $F(t)$ means distribution function. Using equation (24), the hazard function for the Goel-Okumoto model, LM 1 and LM 2 model is defined as follows.

$$h_{Goel-Okumoto}(t) = \beta \, , \; h_{LM1}(t) = \frac{b^2(1+t)}{b+1+bt} \quad ,$$

$$h_{LM2}(t) = \frac{b^2(b+t)}{b+1+bt} \quad . \quad (25)$$

In Figure 4, the Goel-Okumoto model has the same probability of failure at each failure time and the LM 1 and LM 2 models show an increase in failure probability as the failure time passes.

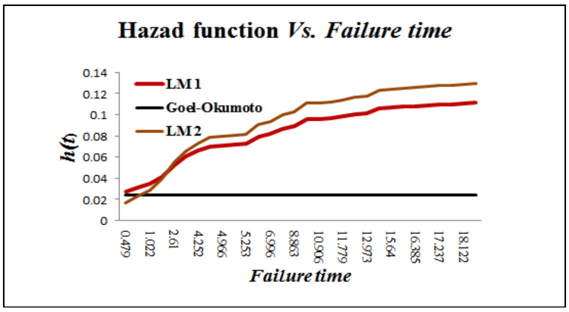

Fig. 4. Trend of hazard function

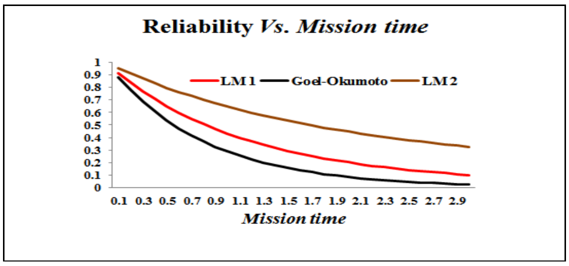

Fig. 5. Transition of reliability pattern

In the NHPP model, a software failure occurs at the time of testing $x_{30} = 18.735$ and reliability which is the probability that a software failure does not occur between 18.735 and $18.735 + t$ (where t is the mission time) can be stated using the ensuing construction [10, 11].

$$\hat{R}(t\,|\,18.735) = e^{-\int_{18.735}^{18.735+t}\lambda(\eta)\,d\eta} = \exp\left[-\left\{m(t+18.735) - m(18.735)\right\}\right] \tag{26}$$

In the form of the reliability function of Figure 5 using the equation (26), it gradually appears as a non-increasing pattern as the mission time pass. Therefore, the LM 1 and LM 2 models are relatively higher than the Goel-Okumoto model in terms of reliability.

IV. CONCLUSION

In the process of software development, it is possible to evaluate the efficiency by comparison and analysis about the software safety by quantitatively the characteristics of the failure or the occurrence of the failure during the execution of the test or actual software operation. This study was compared the reliability characteristics of the software reliability model based on the non-homogeneous Poisson process with finite number of failures using Goel-Okumoto model and the modified Lindley distribution developed to analyze the failure time data. The results of this study are as follows. First, the Goel-Okumoto model in terms of mean squared error appears to be smaller than the Lindley distribution and the modified Lindely distribution, so that the Goel-Okumoto model can be regarded as an efficient model. Secondly, Goel-Okumoto model can be regarded as an efficient model because Goel-Okumoto model than Lindley type model of the estimated value of coefficient for determination has a larger estimated value. However, since the proposed model also has 95% or more, all models are considered to be efficient models. Third, the mean value function patterns of Goel-Okumoto, Lindley and modified Lindley distribution models are similar to each other but the Goel-Okumoto model has a relatively smaller value than the Lindley distribution and the modified Lindley distribution model. Fourth, the hazard function which means the instantaneous failure rate for the failure time is the same for the Goel-Okumoto model at each failure time and the Lindley distribution and the modified distribution model have increase form as the failure time passes. Fifth, in the form of the reliability function, it gradually appears as a non-increasing pattern as the mission time pass. But, the Lindley distribution and the modified Lindley distribution models are relatively higher than the Goel-Okumoto model in terms of reliability. Through this study, software operators can use the mean squared error, mean value function, hazard function, and reliability trend to identify the type of failure in software reliability that reflects various life time distribution characteristics. Using content of this study, it can be concluded that the software design segment can be assisted the software design by using the software failure analysis and applying several life distributions.

Acknowledgements

Funding for this paper was provided by Namseoul University

REFERENCES

[1] K. Y. Song, I. H. Chang, H. Pham, A Software Reliability Model with a Weibull Fault Detection Rate Function Subject to Operating Environments. Applied Science. 2017; 7(983):1-16.

[2] S. Yamada, S. Osaki, Software reliability growth modeling: models and applications. IEEE Transactions on Software Engineering. 1985; 11(12): 1431-1437.

[3] X. Teng, H. Pham, A new methodology for predicting software reliability in the random field environments. IEEE Transactions on Reliability. 2006; 55(3):458-468.

[4] Hee-Cheul Kim, The Comparison Analysis about Reliability Features of Software Reliability Model Using Burr-XII and Type-2 Gumbel Lifetime Distribution. International Journal of Engineering Research and Technology. 2919; 12(1):73-78.

[5] Goel, A. L, Okumoto K, Time-dependent fault detection rate model for software and other performance measures. IEEE Trans. Reliability. 1978; 28:206-211.

[6] Al. Turk, L. I, Characteristics and Application of the NHPP Log-Logistic Reliability Model. International Journal of Statistics and Probability. 2019;8(1):44-55.

[7] S. Karuppusamy, V. Balakrishnan and K. Sadasivan, Modified one-Parameter Lindley Distribution and Its Applications. International Journal of Engineering Research and Applications. 2018; 8(1): 50-56.

[8] Shanker, R., Shanker distribution and Its Applications. International Journal of Statistics and Applications. 2015; 5(6):338-348.

[9] Y. Hayakawa, G. Telfar, Mixed Poisson-type processes with application in software reliability. Mathematical and Computer Modelling. 2000: 151-156.

[10] Hee-Cheul Kim. A Comparative Study on Software Reliability Models with Shape Parameter of Type-2 Gumble Life Distribution. International Journal of Soft Computing. 2017; 12(5-6):351-354.

[11] Kim H-C, Kim K-S, Software Development Cost Model based on NHPP Gompertz Distribution. Indian Journal of Science and Technology.2016; 9(20):1-6.

[12] Kim K-S, Kim H-C, The Performance Analysis of the Software Reliability NHPP Log-linear Model Depend on Viewpoint of the Learning Effects, Indian Journal of Science and Technology. 2016; 9(37):1-5.

[13] https://www.r-project.org/

[14] Kuei-Chen, C., Yeu-Shiang, H., and Tzai-Zang, L., A study of software reliability growth from the perspective of learning effects, Reliability Engineering and System Safety 93. 2008:1410-1421.