

# 유한고장 NHPP 어랑분포의 형상모수 변화에 따른 소프트웨어 신뢰성 모형의 속성 분석에 관한 연구

민 경 일\*

## A Study on the Property Analysis of Software Reliability Model with Shape Parameter Change of Finite Fault NHPP Erlang Distribution

Kyung Il Min\*

### Abstract

Software reliability has the greatest impact on computer system reliability and software quality. For this software reliability analysis, In this study, we compare and analyze the trends of the properties affecting the reliability according to the shape parameters of Erlang distribution based on the finite fault NHPP. Software failure time data were used to analyze software failure phenomena, the maximum likelihood estimation method was used for parameter estimation. As a result, it can be seen that the intensity function is effective because it shows a tendency to decrease with time when the shape parameters  $a = 1$  and  $a = 3$ . However, the pattern of the mean value function showed an underestimation pattern for the true values when the shape parameters  $a = 1$  and  $a = 2$ , but it was found to be more efficient when  $a = 3$  because the error width from the true value was small. Also, in the reliability evaluation of the future mission time, the stable and high trend was shown when the shape parameters  $a = 1$  and  $a = 3$ , but on the contrary, when  $a = 2$ , the reliability decreased with the failure time. Through this study, the property of finite fault NHPP Erlang model according to the change of shape parameter without existing research case was newly analyzed, and new research information that software developers can use as basic guideline was presented.

Keywords : Infinite Failure, NHPP, Shape Parameter, Software Failure Reliability Model, Erlang  
Distribution Property Analysis

## 1. 서 론

소프트웨어 관련 산업이 급속하게 성장함에 따라 대용량의 소프트웨어를 사용하는 응용 분야도 함께 크게 증가함으로써 소프트웨어 신뢰성 평가에 관한 어려움도 점차 증가하고 있다. 이러한 소프트웨어 산업의 생태계 변화로 소프트웨어 품질을 향상시키기 위한 소프트웨어 신뢰성 평가기술이 더욱 절실히 필요해지고 있다.

소프트웨어를 최종 출시하기 전 사전 테스트를 통하여 소프트웨어의 시스템 요구사항을 충족시키는지를 점검하게 된다. 이 문제는 사용자 입장에서의 요구조건 및 효율성, 비용 등을 만족시켜야 한다. 소프트웨어 테스트 비용을 줄이기 위해서는 사전에 소프트웨어 신뢰성의 속성과 비용을 알고 있어야 효율적이다. 즉, 소프트웨어 제품의 고장 내용을 예측하기 위한 신뢰성 모형 개발이 절실하여, 지금까지 많은 신뢰성 모형이 제안 되었다. 이 중에서 비동질적 포아송 과정(Non-homogeneous Poisson process; NHPP)을 적용한 모형[Gokhale and Trivedi, 1999]은 여러 탐색 과정측면에서 효율적인 모형이고, 결함(고장)이 발생하면 즉시, 제거되며 오류를 수정하는 과정에서 새로운 결함이 발생되지 않는다는 가정을 하고 있다. 이 분야에서 Gokhale and Trivedi[1999]은 개선된 비동질적인 포아송 과정(Enhanced NHPP)을 제시하였고 Goel and Okumoto[1978]은 지수적인 소프트웨어 신뢰성 모형(exponential software reliability growth model)을 제안 하였다. 또한, Yang[2016]은 Lomax 분포의 형상모수 변화에 따른 소프트웨어 개발비용 모형에 관한 연구를 하기도 하였다. 하지만, 기존 연구들은 효율적인 신뢰성 예측 방법과 신뢰성 검사도구에 대한 학습과정을 제시하는 내용이 대부분이며, 고장분포의 형태를 결정하는 형상모수에 관한 연구가 없었다.

따라서, 본 연구에서는 유한고장 NHPP 모형에 근

거하여, 기존 연구사례가 없는 어랑분포(Erlang 형상 모수 변화에 따른 소프트웨어 신뢰성 모형의 속성을 비교, 분석하고자 한다.

## 2. NHPP 신뢰성 모형

$N(t)$  을 시간  $t$  까지 탐지된 소프트웨어의 고장의 누적수이고  $m(t)$  를 이에 대한 기대값을 나타내는 평균값 함수,  $\lambda(t)$  을 강도함수(Intensity function) 로 표시하면, 누적 고장수  $N(t)$  는 모수  $m(t)$  을 가진 포아송 확률밀도함수(Probability density function)를 따른다고 알려져 있다[Gokhale and Trivedi, 1999; Yoo, 2015]. 즉,

$$P\{N(t) = n\} = \frac{[m(t)]^n \cdot e^{-m(t)}}{n!} \quad (1)$$

단,  $\infty$

NHPP 모형의 평균값 함수  $m(t)$  와 강도함수  $\lambda(t)$  는 다음과 같다[Kim, 2016].

$$m(t) = \int_0^t \lambda(s) ds \quad (2)$$

$$\frac{dm(t)}{dt} = \lambda(t) \quad (3)$$

본 연구에서 제시하는 유한고장 NHPP 모형에서는 시간  $(0, t]$  까지 탐지되어 질 수 있는 결함의 기대값을  $\theta$  로 나타내면, 평균값 함수와 강도함수는 다음과 같다[Gokhale and Trivedi, 1999; Kim, 2015].

$$m(t) = \theta F(t) \quad , \quad \lambda(t) = \theta F'(t) \quad (4)$$

여기서,  $\theta$  을 모수공간이라 하면 NHPP 모형의 우도함수는 다음과 같다[Kim, 2015].

$$L_{NHPP}(\theta | \underline{x}) = \left( \prod_{i=1}^n \lambda(x_i) \right) \exp[-m(x_n)] \quad (5)$$

단,  $\underline{x} = (x_1, x_2, x_3, \dots, x_n)$

또한, NHP 모형의 테스트 시점  $x_n$ 에서 소프트웨어 고장이 발생한다는 가정하에 신뢰구간  $[x_n, x_n + \tau]$  (단,  $\tau$ 는 임무시간(Mission time)사이에서 소프트웨어 고장이 일어나지 않을 확률인 신뢰도

$\hat{R}(t | x_n)$ 는 다음과 같다[Yoo, 2015].

$$\hat{R}(t | x_n) = e^{-\int_{x_n}^{x_n+t} \lambda(\tau) d\tau} \quad (6)$$

$$= \exp[-\{m(t + x_n) - m(x_n)\}]$$

### 3. 유한고장 NHPP 어랑분포

어랑분포는 신뢰성에서 많이 사용하는 분포로서 다양한 현상을 나타낼 수 있으며, 형상모수 ( $a$ )와 척도모수( $b$ )에 따른 확률밀도함수와 누적분포는 다음과 같다[Yoo, 2015].

$$f(t) = \frac{b^a}{\Gamma(a)} t^{a-1} e^{-bt} \quad (7)$$

$$F(t) = \left(1 - e^{-bt} \sum_{i=0}^{a-1} \frac{(bt)^i}{i!}\right) \quad (8)$$

여기서,  $a, b > 0, a = 1, 2, 3 \dots, t \in [0, \infty]$

관측시간  $(0, t]$ 까지 검출하기 위한 시간절단 (Time truncated) 모형은  $n$  번째까지 고장시점 자료  $x_n$ 은 다음과 같은 관계로 표시된다[Rohatgi, 1984].

$$x_n = \sum_{i=1}^n t_i \quad (9)$$

$$(i = 1, 2, \dots, n; 0 \leq x_1 \leq x_2 \leq \dots \leq x_n)$$

이 경우에 확률밀도함수  $f(\bullet)$  와 누적분포  $F(\bullet)$ 을 이용하면 NHPP 모형의 우도함수는 다음과 같다[Kim, 2016].

$$L_{NHPP}(\Theta | \underline{x}) = \left(\prod_{i=1}^n \theta f(x_i)\right) \exp[-\theta F(x_n)] \quad (10)$$

단,  $\underline{x} = (x_1, x_2, x_3, \dots, x_n)$ ,  $\Theta$ 은 모수공간을 의미한다.

따라서 식 (10)을 이용하면 유한고장 NHPP 모형의 로그우도함수는 다음과 같다[Rao et al., 2011].

$$\ln L_{NHPP}(\Theta | \underline{x}) = \quad (11)$$

$$n \ln \theta - n \ln \Gamma(a) + na \ln b + (a-1) \sum_{i=1}^n \ln x_i$$

$$- b \sum_{i=1}^n x_i - \theta + \theta e^{-b x_n} \left( \sum_{i=0}^{a-1} \frac{(b x_n)^i}{i!} \right)$$

식 (11)에서  $\theta$  와  $b$ 에 대하여 편미분하면 최우 추정값  $\hat{\theta}_{MLE}$ 와  $\hat{b}_{MLE}$ 을 수치 해석으로 추정 할 수 있다.

$$\frac{\partial \ln L_{NHPP}(\Theta | \underline{x})}{\partial \theta} = \quad (12)$$

$$\frac{n}{\theta} - 1 + e^{-b x_n} \left( \sum_{i=0}^{a-1} \frac{(b x_n)^i}{i!} \right) = 0$$

$$\frac{\partial \ln L_{NHPP}(\Theta | \underline{x})}{\partial b} = \quad (13)$$

$$\frac{a n}{b} - \sum_{i=1}^n x_i + \frac{\partial \left[ \theta e^{-b x_n} \left( \sum_{i=0}^{a-1} \frac{(b x_n)^i}{i!} \right) \right]}{\partial b} = 0$$

본 연구에서는 고장분포의 형태를 결정하는 형상 모수의 변화에 따라서 유한고장 NHPP 모형의 속성의 변화를 비교, 분석하기 위하여, 형수 모수가 각각  $a = 1, 2, 3$ 인 경우를 고려한다[Yang, 2016; Kim, 2018].

가. 형상모수  $a = 1$ 인 경우의 최우추정값

$$\frac{\partial \ln L_{NHPP}(\Theta | \underline{x})}{\partial \theta} = \frac{n}{\theta} - 1 + e^{-b x_n} = 0 \quad (14)$$

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

나. 형상모수  $a = 2$  인 경우의 최우추정값

$$\frac{\partial \ln L_{NHPP}(\Theta | \underline{x})}{\partial \theta} = \frac{n}{\theta} - 1 + e^{-bx_n} (1 + bx_n) = 0 \tag{16}$$

$$\frac{\partial \ln L_{NHPP}(\Theta | \underline{x})}{\partial b} = \frac{2n}{b} - \theta b x_n^2 e^{-bx_n} - \sum_{i=1}^n x_i = 0 \tag{17}$$

다. 형상모수  $a = 3$  인 경우의 최우추정값

$$\frac{\partial \ln L_{NHPP}(\Theta | \underline{x})}{\partial \theta} = \frac{n}{\theta} - 1 + e^{-bx_n} \left( 1 + bx_n \frac{x_n^2}{2} b^2 \right) = 0 \tag{18}$$

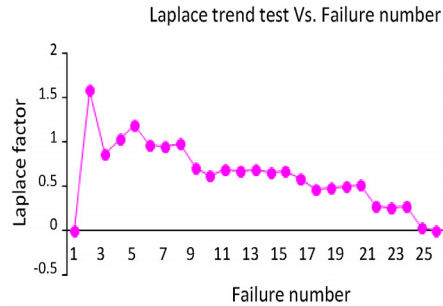
$$\frac{\partial \ln L_{NHPP}(\Theta | \underline{x})}{\partial b} = \frac{3n}{b} - \theta b^2 \frac{x_n^3}{2} e^{-bx_n} - \sum_{i=1}^n x_i = 0 \tag{19}$$

#### 4. 소프트웨어 고장시간 분석

이 절에서는 <Table 1>과 같은 소프트웨어 고장 시간자료[Rao et al., 2011](Failure time data)를 사용하여 신뢰도 성능을 분석하고자 한다. 이를 위해 먼저, 자료에 대한 신뢰성을 확보하고자 라플라스 추세검정(Laplace trend test)을 실시하였다[Kim, 2015]. 그 결과는 <Figure 1>에 요약되었다. 이 그림에서 라플라스 요인(Factor)이 “-2와 +2” 사이에 존재함으로써 이 자료를 이용하여 신뢰성 모형을 제시하는 것이 가능하다는 것을 나타내고 있다[Kim, 2016, 2015].

<Table 1> Software Failure Time Data

Failure number	Failure time (hours)	Failure numbe	Failure time(hours)
1	30.02	16	151.78
2	31.46	17	177.50
3	53.93	18	180.29
4	55.290	19	182.21
5	58.720	20	186.34
6	71.920	21	256.81
7	77.070	22	273.88
8	80.900	23	277.87
9	101.90	24	453.93
10	114.87	25	535.00
11	115.34	26	537.27
12	121.57	27	552.9
13	124.97	28	673.68
14	134.07	29	704.49
15	136.25	30	738.68



<Figure 1> Laplace Trend Test

또한, 모수추정은 최우추정법을 이용하였고, 수치 해석 방법인 이분법(Bisection method)을 사용하여 비선형 방정식을 계산하였다. 이러한 계산은 C-언어 프로그래밍을 이용하였고, 초기값을 0.001과 5, 허용한계(Tolerance for width of interval)는  $10^{-5}$ 을 주고 수렴성을 확인하면서 모수를 추정하였다.

최우추정법의 결과와  $MSE$ (평균제곱오차)와  $R^2$ (결정계수)가 <Table 2>에 요약되었다.

<Table 2> Parameter Estimation of the Each Model and  $MSE, R^2$

Shape parameter	MLE		Model comparison	
	$\hat{\theta}$	$\hat{b}$	$MSE$	$R^2$
a=1	33.4092	0.3089	45.7466	0.8550
a=2	30.5978	0.0079	45.4879	0.8558
a=3	31.2672	1.2425	49.1920	0.8441

Note)  $MLE$  : Maximum likelihood estimation;  
 $MSE$  : Mean square error;  
 $R^2$  : Coefficient of determination.

<Table 2>에서 평균제곱오차(MSE)는 실제 관찰 값과 예측 값에 대한 차이를 나타내는 측도로서, 모형 비교에 있어서 이 값이 작을수록 상대적으로 우수한 모형이 된다[Yang, 2016; Kim, 2016].

$$MSE = \frac{\sum_{i=1}^n (m(x_i) - \hat{m}(x_i))^2}{n - k} \quad (20)$$

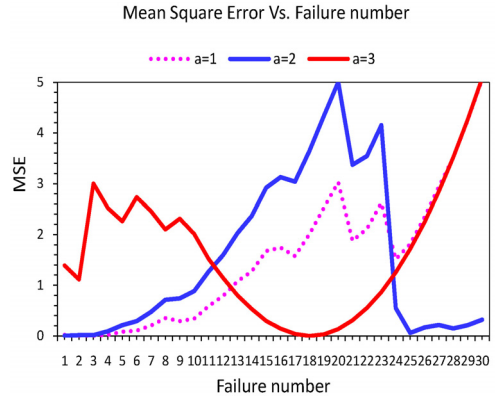
여기서,  $m(x_i)$ 은 시간(0,  $x_i$ ]까지 나타난 에러들의 누적함수이고,  $\hat{m}(x_i)$ 는  $x_i$  시점까지 평균값 함수로부터 추정된 에러의 누적개수를 의미한다. 그리고  $n$ 은 관찰 값의 수,  $k$ 는 모수의 개수를 의미한다.

결정계수( $R^2$ )는 관찰 값의 차이에 대한 예측 값의 설명력을 나타내는 측도로서, 모형 비교에 있어서 결정계수 값이 클수록 상대적으로 우수한 모형이 된다[Kim, 2016].

$$R^2 = 1 - \frac{\sum_{i=1}^n (m(x_i) - \hat{m}(x_i))^2}{\sum_{i=1}^n (m(x_i) - \sum_{j=1}^n m(x_j)/n)^2} \quad (21)$$

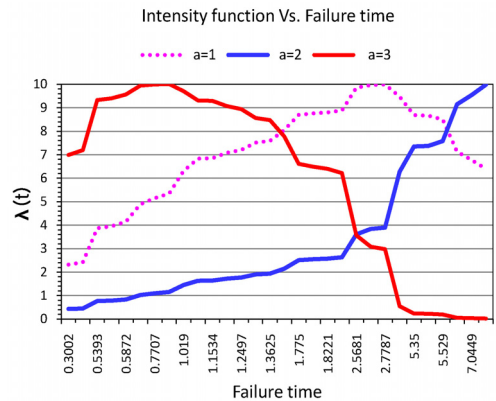
그러므로, <Table 2>에서와 같이 제안한 형상모수

값들 중에서, 형상모수  $a = 2$ 일 때 평균제곱오차가 제일 작고, 결정계수도 크게 나타나기 때문에 가장 효율적이라고 판단할 수 있다. 여기서, <Figure 2>는 평균제곱오차(MSE)의 추이를 나타내고 있다[Teng, and Pham, 2006; Yang, 2017].



<Figure 2> Transition of Mean Square Error(MSE)

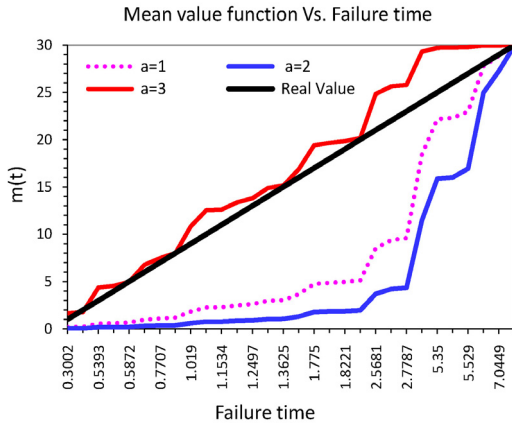
<Figure 3>은 고장시간에 대한 강도함수의 속성이 요약 되었다. 이 그림에서 형상모수  $a=1$ 과  $a = 3$ 일 때 모두 고장시간이 흘러 갈수록 감소하는 추세를 보이고 있으나,  $a = 3$ 일 때가 가장 효율적임을 알 수 있다.



<Figure 3> Transition of Intensity Function  $\lambda(t)$

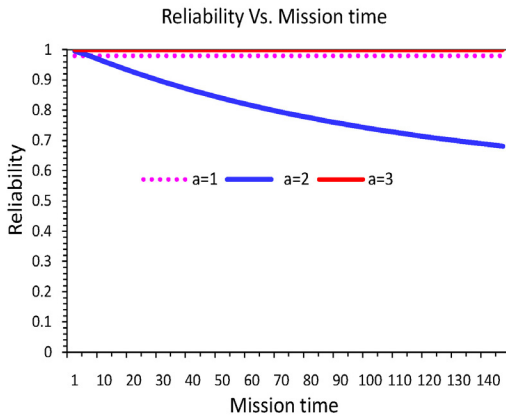
<Figure 4>는 평균값 함수에 대한 속성이 요

약 되었다.이 그림에서 형상모수  $a=3$ 일 때 참값에 대해 약간의 과대평가 패턴을 보이고 있으나, 크게 과소평가 패턴을 보이는 형상모수  $a=1$ 과  $a=2$ 보다는 오차의 폭이 작아 효율적임을 알 수 있다.



<Figure 4> Pattern of Mean Value Function  $m(t)$

<Figure 5>는 미래의 임무시간에 대한 신뢰도 속성을 보여 주고 있다. 즉, 형상모수  $a=2$ 일 때는 임무시간이 흘러갈수록 신뢰도는 감소하지만, 형상모수  $a=1$ 과  $a=3$ 일 때는 임무시간이 흘러도 안정적으로 높은 신뢰도 추세를 보이고 있음을 알 수 있다[Yang, 2017].



<Figure 5> Transition of Reliability  $\hat{R}(t)$

## 5. 결 론

소프트웨어 테스트링 작업이나, 실제 소프트웨어 개발 과정에서 고장발생에 대한 속성을 정량적으로 모형화한 후 특성을 분석하면 상대적으로 효율적인 평가를 할 수 있다. 따라서, 본 연구에서는 효율적인 평가를 위해 기존 연구사례가 없는 유한고장 NHPP 어랑모형의 형상모수 변화에 따른 소프트웨어 신뢰속성의 추이를 비교 분석하였다.

본 연구의 결과는 다음과 같다.

첫째, 강도함수 속성 측면에서는 형상모수  $a=1$ 과  $a=3$ 일 때 모두 고장시간이 흘러 갈수록 감소하는 추세를 보여서 효율적임을 알 수 있으나, 반대로  $a=2$ 일 때는 증가함을 알 수 있었다.

둘째, 평균값 함수  $m(t)$ 의 패턴 분석에 있어서 형상모수  $a=1$ 과  $a=2$ 일 때 참값에 대해 오차폭이 크고, 크게 과소 평가 패턴을 보이고 있으나, 형상모수  $a=3$ 일 때 참값과의 오차의 폭이 작아 가장 우수함을 알 수 있었다.

셋째, 신뢰도에 있어서는 형상모수  $a=1$ 과  $a=3$ 일 때 안정적으로 높은 신뢰도 추이(Transition)를 보이고 있으나, 반대로  $a=2$ 일 때는 고장시간이 흐를수록 신뢰도가 감소하는 것을 알 수 있었다. 즉, 본 연구를 통해 전반적으로 형상모수  $a$ 값이 클수록 효율적임을 확인할 수 있었다.

결론적으로, 본 연구결과를 통하여 기존 연구 사례가 없는 유한고장 NHPP Erlang 모형의 형상모수의 변화에 따른 속성 추이에 대한 새로운 분석과 함께, 소프트웨어 개발자들이 기본적인 설계지침으로 활용할 수 있는 사전 정보를 제공할 수 있었다. 더불어, 소프트웨어 개발자들에게 필요한 정보를 제공하기 위하여 추후에 같은 형태의 고장분포를 갖는 다른 모형들과의 신뢰성 비교, 분석을 통해 최적의 신뢰성 모형을 찾아내는 후속 연구가 필요하다.

## References

- [1] Goel, A. L. and Okumoto, K., "Time-dependent fault detection rate model for software and other performance measures", *IEEE Trans. Reliab.*, Vol. 28, 1978, pp. 206-211.
- [2] Gokhale, S. S. and Trivedi, K. S. A., "time/structure based software reliability model", *Annals of Software Engineering*, Vol. 8, 1999, pp. 85-121.
- [3] Kim, H. C., "A Performance Analysis of Software Reliability Model using Lomax and Gompertz Distribution Property", *Indian Journal of Science and Technology*, Vol. 9, No. 20, 2016, pp. 1-6.
- [4] Kim, H. C., "The Property of Learning effect based on Delayed Software S-Shaped Reliability Model using Finite NHPP Software Cost Model", *Indian Journal of Science and Technology*, Vol. 8, No. 34, 2015, pp. 1-7.
- [5] Rao, K. H., Prasad, R. S., and Kantham, R. L., "Software Reliability Measuring using Modified Maximum Likelihood Estimation and SPC", *International Journal of Computer Applications*, Vol. 21, No. 7, May 2011, pp. 1-5.
- [6] Rohatgi, V. K., *Statistical inference*, John Wiley & Sons, Inc, New York, 1984, pp. 398-416.
- [7] Teng, X. and Pham, H., "A new methodology for predicting software reliability in the random field environments", *IEEE Transactions on Reliability*, Vol. 55, No. 3, 2006, pp. 458-468.
- [8] Yang, T. J., "A Software Reliability Cost Mode Based on the Shape Parameter of Lomax Distribution", *The Journal of Korea Institute of Information Electronics and Communication Technology*, Vol 9, No 2, 2016, pp. 171-177.
- [9] Yang, T. J., "The Performance Analysis Comparative Study depend on Software Reliability Model and Curve Regression Model", *Medwell Journals*, Vol. 12, No. 5, 2017, pp. 313-317.
- [10] Yoo, T. H., "The Infinite NHPP Software Reliability Model based on Monotonic Intensity Function", *Indian Journal of Science and Technology*, Vol. 8, No. 14, 2015, pp. 1-7.

## ■ 저자소개



Kyung Il Min

He received a MS and a Ph.D from the Department of Management Information Engineering at Inha University and Seokyeong University, respectively. He served as General Director of human resources development service of Korea. Currently, he has been a professor in the Industry Academic Cooperation Foundation at Namseoul University since 2017. His current research interests include information communication policy, computer system analysis.