

어랑 NHPP 소프트웨어 모형에 근거한 최적방출시기 비교 연구

김희철*

요약

본 연구에서는 소프트웨어 제품을 개발하여 테스트를 거친 후 사용자에게 인도하는 시기를 결정하는 방출문제에 대하여 연구 하였다. 소프트웨어의 결함을 제거하거나 수정 작업 중에도 새로운 결함이 발생될 가능성이 있는 경우 무한고장수를 가진 비동질적인 포아송 과정에 기초하였다. 수명강도는 다양한 형상모수와 척도모수에 이용 할 수 있기 때문에 신뢰성 분야에서 많이 사용되는 어랑 강도분포를 이용한 방출시기에 관한 문제를 제시하였다. 소프트웨어 요구 신뢰도를 만족시키고 소프트웨어 개발 및 유지 총비용을 최소화 시키는 최적 소프트웨어 방출 정책에 대하여 논의 되었다. 본 논문의 수치적인 예에서는 고장 간격 시간 자료를 적용하였으며 모수추정 방법은 최우추정법을 이용하여 추세분석으로 자료의 효율성을 입증한 후 최적 방출시기를 추정하였다.

The Comparative Study of Optimal Release Time Based on Erlang Non-Homogeneous Poisson Process Software Model

Hee-Cheul Kim*

ABSTRACT

In this paper, make a study decision problem called an optimal release policies after testing a software system in development phase and transfer it to the user. When correcting or modifying the software, because of the possibility of introducing new faults when correcting or modifying the software, infinite failure non-homogeneous Poisson process models presented and propose release policies of the life distribution applied Erlang distribution model which used to an area of reliability because of various shape and scale parameter. In this paper, discuss optimal software release policies which minimize a total average software cost of development and maintenance under the constraint of satisfying a software reliability requirement. In a numerical example, after trend test applied and estimated the parameters using maximum likelihood estimation of inter-failure time data, make out estimating software optimal release time.

Key Words : Software Release Policies, NHPP, Erlang Distribution Function, Reliability Growth, Mission Time.

* 남서울대학교 산업경영공학과

· 제1저자(First Author) : 김희철 · 교신저자(Correspondent Author) : 김희철
· 접수일(2010년 2월 16일), 수정일(1차 : 2010년 3월 30일), 게재확정일(2010년 4월 2일)

1. 서 론

소프트웨어 방출시간에 대한 연구들은 대부분 유한 고장 NHPP(Non-Homogeneous Poisson Process) 모형을 사용하였다[1, 2].

이러한 유한(Finite)고장 NHPP모형은 소프트웨어가 유한개의 고장이 있고 고장 제거 단계에서는 새로운 고장이 발생하지 않는다는 가정을 한 모형이다. 그러나 실제 고장 제거 단계에서도 새로운 고장이 발생할 수 있다. 따라서 본 연구에서는 무한(Infinitive) 고장 NHPP 모형을 이용하여 최적 방출시기에 대한 문제를 제안하고자 한다. 이 분야에서는 Musa-Okumoto의 대수 포아송 실행시간 모형[3]과 로그-파우어 모형[4, 5, 12]을 이용한 방출 문제에 대한 문제들이 이미 연구되었고 최근까지도 이와 관련된 연구는 Yang과 Xie[13], Huang 등에 의해 보증기간을 포함한 방출시기 등에 대하여 연구되고 있다. 본 연구에서는 소프트웨어의 결함을 제거하거나 수정 작업 중에도 새로운 결함이 발생할 가능성이 있는 무한고장수를 가진 비동질적인 포아송 과정에 기초하고 수명분포는 다양한 형상모수와 척도모수에 이용 할 수 있기 때문에 신뢰성 분야에서 많이 사용되는 어랑형 분포함수를 이용한 최적 방출시기에 관한 문제를 다루었다.

II. 관련 연구

2.1 무한고장 NHPP와 기록값 통계량

NHPP 모형에서 평균값 함수 $m(t)$ (Mean value function)와 강도 함수(Intensity function) $\lambda(t)$ 는 다음과 같은 관계로 표현할 수 있다.

$$m(t) = \int_0^t \lambda(s) ds, \frac{dm(t)}{dt} = \lambda(t) \quad (1)$$

따라서 $N(t)$ 는 모수 $m(t)$ 을 가진 포아송 확률 밀도 함수(Probability density function)로 알려져 있다. 즉,

$$P\{N(t) = n\} = \frac{[m(t)]^n \cdot e^{-m(t)}}{n!}, n = 0, 1, 2, \dots, \infty \quad (2)$$

이처럼 시간 관련 모형(Time domain models) 들은 NHPP에 의해서 확률 고장 과정으로 설명이 가능하다. 이러한 모형들은 고장 강도 함수 $\lambda(t)$ 가 다르게 표현됨으로서 평균값 함수 $m(t)$ 도 역시 다르게 나타나고 이러한 NHPP 모형들은 유한 고장 모형과 무한 고장 범주로 분류한다.

우리 시점에서 고장이 발생할 수도 있는 상황을 반영하기 위하여 RVS(Record Value Statistics)모형을 사용하는 NHPP 모형의 평균값함수는 다음과 같이 알려져 있다.

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

따라서 (1)식 과 (3)식을 연관시키고 $f(t)$ 을 확률밀도함수, $F(t)$ 을 분포함수라고 하면 다음과 같은 관계식에 의해 NHPP의 강도함수는 $F(t)$ 의 위험함수($h(t)$)가 된다.

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

시간 $(0, t]$ 까지 조사하기 위한 시간 절단(Time truncated) 모형은 n 번째 까지 고장 시점 자료를

$$x_k = \sum_{i=1}^k t_k \quad (k = 1, 2, \dots, n; 0 \leq x_1 \leq x_2 \leq \dots \leq x_n) \quad (5)$$

이라고 하고 θ 을 모수공간이라고 하면 우도 함수는 다음과 같이 알려져 있다.

$$L_{NHPP_{FNV}}(\theta | D_t) = \prod_{i=1}^n \left(\frac{f(x_i)}{1-F(x_i)} \right) (1-F(x_n)) \quad (6)$$

$$= \left(\prod_{i=1}^n h(x_i) \right) (1-F(x_n))$$

$$\frac{\partial \ln L(\phi | \underline{x})}{\partial \phi} = \quad (11)$$

$$\frac{n}{\phi} - \sum_{i=1}^n \frac{x_i}{\phi x_i + 1} - \frac{n x_n}{(\phi x_n + 1) \ln(\phi x_n + 1)} = 0$$

2.2 기존모형인 로그 포아송 실행시간모형 및 로그 파워어 모형

로그 포아송 실행시간(Log Poission execution) 모형[5, 11]은 1984년에 Musa 와 Okumoto에 의해서 소개된 무한 고장 소프트웨어 모형으로 평균값함수와 강도함수는 다음과 같이 알려져 있다.

$$m(t) = \frac{1}{\theta} \ln(\lambda_0 \theta t + 1), \lambda(t) = \frac{\lambda_0}{\lambda_0 \theta t + 1} \quad (7)$$

한편, (7)식을 (6)식에 대입하면 우도함수는 다음과 같이 유도 할 수 있다.

$$L(\theta, \lambda_0 | \underline{x}) = \left(\prod_{i=1}^n \frac{\lambda_0}{\lambda_0 \theta x_i + 1} \right) \cdot e^{-\frac{1}{\theta} \ln(\lambda_0 \theta x_n + 1)} \quad (8)$$

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

최우추정법(MLE)을 이용하기 위한 로그 우도 함수는 다음과 같이 유도된다.

$$\ln L(\theta, \lambda_0 | \underline{x}) \quad (9)$$

$$= n \ln \lambda_0 - \ln \prod_{i=1}^n (\lambda_0 \theta x_i + 1) - \frac{1}{\theta} \ln(\lambda_0 \theta x_n + 1)$$

(9)식을 이용하여 최우추정치 $\hat{\theta}_{MLE}$ 와 $\hat{\lambda}_{0MLE}$ 은 다음과 같이 구할 수 있다고 하였다.

$$\hat{\theta}_{MLE} = \frac{1}{n} \ln(\hat{\phi} x_n + 1), \hat{\lambda}_{0MLE} = \hat{\phi} / \hat{\theta}_{MLE} \quad (10)$$

$\hat{\phi} (= \hat{\lambda}_{0MLE} \cdot \hat{\theta}_{MLE})$ 는 (11) 식의 근이 되고 이 $\hat{\phi}$ 근을 구하기 위해서는 수치 해석적 방법으로 다음과 같은 식을 이용하여 계산할 수 있다.

로그 파워어 (Log Power)모형[5, 12]은 1999년에 Xie와 Homg에 의해서 발전된 무한 고장 소프트웨어 모형으로 평균값함수와 강도함수는 다음과 같이 알려져 있다.

$$m(t) = a \ln^b(1+t), \lambda(t) = \frac{a b \ln^{b-1}(1+t)}{1+t} \quad (12)$$

한편, (6)식과 (12)식을 이용하면 우도함수는 다음과 같이 유도 할 수 있다.

$$L(a, b | \underline{x}) = \left(\prod_{i=1}^n \frac{a b \ln^{b-1}(1+x_i)}{1+x_i} \right) \cdot e^{-a \ln^b(1+x_n)} \quad (13)$$

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

최우추정법(MLE)을 이용하기 위한 로그 우도 함수는 다음과 같이 유도된다.

$$\ln L(a, b | \underline{x}) = n \ln a + n \ln b - (b-1) \ln \left(\sum_{i=1}^n \ln(1+x_i) \right) - \sum_{i=1}^n \ln(1+x_i) - a \ln^b(1+x_n) \quad (14)$$

(14)식에서 a 와 b 에 대하여 편미분 하여 다음과 같은 식을 만족하는 \hat{a}_{MLE} 와 \hat{b}_{MLE} 을 수치 해석적 방법으로 계산할 수 있다.

$$\frac{\partial \ln L(a, b | \underline{x})}{\partial a} = \frac{n}{a} - \ln^b(1+x_n) = 0 \quad (15)$$

$$\frac{\partial \ln L(a, b | \underline{x})}{\partial b} = \frac{n}{b} - \ln \left(\sum_{i=1}^n \ln(1+x_i) \right) - a \ln^b(1+x_n) \ln(\ln(1+x_n)) = 0 \quad (16)$$

2.3 어랑분포에 기초한 NHPP모형

소프트웨어 신뢰성 분야에서 널리 사용되는 분포 중에서 감마분포는 형상(shape) 및 척도(scale)모수의 값에 따라 다양한 분포를 표현할 수 있어 신뢰성 데이터 분석에 가장 널리 사용된다. 감마 분포의 확률밀도 함수와 분포함수는 다음과 같이 정의 된다.

$$f(t) = \frac{\beta(\beta t)^{\alpha-1}}{\Gamma(\alpha)} \exp(-\beta t), \quad \alpha, \beta, t > 0 \quad (17)$$

$$F(t) = 1 - R(t) = 1 - \sum_{k=0}^{\alpha-1} \frac{(\beta t)^k}{k!} \exp(-\beta t) \quad (18)$$

위 식에서 α 는 형상모수이고 β 는 척도 모수이다. 형상모수 α 는 이 분포의 비대칭성(왜도, skewness), 첨도(kurtosis) 등 분포의 형상을 결정하는 모수로 중요한 위치를 차지하고 있다. α 에 따른 감마분포의 신뢰도와 고장률은 그림 1에 나타내었다.

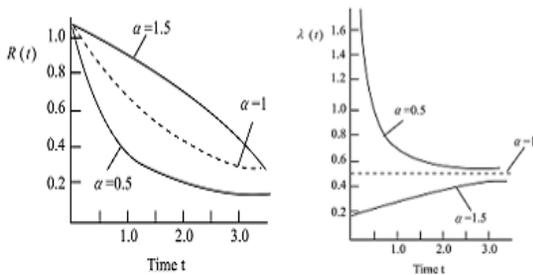


그림 1. 감마분포의 신뢰도와 고장률($\beta = 0.5$)
 Fig. 1. Reliability and failure rate of gamma distribution

이러한 감마분포는 $\alpha = 1$ 일 때 확률 밀도 함수는 모수 β 을 갖는 지수 분포와 동일하고 고장률은 상수가 된다. 즉, (18)식의 분포함수는 폐쇄형(Close form)이기 때문에 $\alpha = n$ 일 때(n 은 정수)일 때 어랑 분포로 알려져 있다[3, 10]. 따라서 본 논문에서 사용하고자는 어랑분포(Erlang distribution)모형은 강도함수가 증가 패턴을 가지는 $\alpha = 2$ 와 $\alpha = 3$ 인 특수한 경우를 사

용하고자 한다. $\alpha = 2$ 인 경우의 확률밀도함수는 다음과 같다.

$$f(t) = \beta^2 t \exp(-\beta t) \quad (19)$$

따라서 식(1)과 식(3)을 연관 시킨 무한 고장 NHPP 모형에 적용시킨 어랑분포의 강도 함수와 평균값 함수는 다음과 같다.

$$\lambda(t) = \beta^2 t / (\beta t + 1), \quad m(t) = \beta t - \ln(1 + \beta t) \quad (20)$$

어랑분포 모형에 대한 최우추정법(MLE)을 이용하기 위한 로그 우도 함수는 다음과 같이 유도된다.

$$\begin{aligned} \ln L(2, \beta | \underline{x}) &= 2n \ln \beta - \sum_{i=1}^n \ln(\beta x_i + 1) + \sum_{i=1}^n \ln x_i - \beta x_n + \ln(1 + \beta x_n) \end{aligned} \quad (21)$$

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

(21)식을 이용하여 β 에 대하여 미분하고 그 값을 0을 만족하는 최우추정치 $\hat{\beta}_{MLE}$ 을 다음 식에서 추정할 수 있다.

$$\frac{2n}{\beta} = \sum_{i=1}^n \frac{x_i}{1 + \beta x_i} + x_n - \frac{x_n}{1 + \beta x_n} \quad (22)$$

$\alpha = 3$ 인 경우에는 $\alpha = 2$ 인 경우와 유사하게 적용시킨 평균값 함수와 강도함수는 다음과 같다.

$$\lambda(t) = \left(\frac{1}{2} \beta^3 t^2 \right) / \left(1 + \beta t + \frac{\beta^2 t^2}{2} \right) \quad (23)$$

$$m(t) = \beta t - \ln \left(1 + \beta t + \frac{\beta^2 t^2}{2} \right) \quad (24)$$

$\alpha=3$ 인 경우에도 최우추정법(MLE)을 이용하기 위한 로그 우도 함수는 다음과 같이 유도된다.

$$\begin{aligned} \ln L(3, \beta | \underline{x}) &= -n \ln 2 + 3n \ln \beta - \sum_{i=1}^n \ln \left(1 + \beta x_i + \frac{\beta^2 x_i^2}{2} \right) \\ &\quad + \sum_{i=1}^n \ln x_i^2 - \beta x_n + \ln \left(1 + \beta x_n + \frac{\beta^2 x_n^2}{2} \right) \end{aligned} \quad (25)$$

(25)식을 이용하여 β 에 대하여 미분하고 그 값을 0으로 하는 최우추정치 $\hat{\beta}_{MLE}$ 을 다음 식을 이용하여 추정할 수 있다.

$$\frac{3n}{\beta} = \sum_{i=1}^n \frac{x_i + x_i^2 \beta}{1 + \beta x_i + \frac{\beta^2 x_i^2}{2}} + x_n - \frac{x_n + x_n^2 \beta}{1 + \beta x_n + \frac{\beta^2 x_n^2}{2}} \quad (26)$$

III. 신뢰도 및 비용을 고려한 방출시간

NHPP 모형에서 테스트 시점 x_n (마지막 고장시점)에서 소프트웨어 고장이 일어난다고 하는 가정 하에서 신뢰구간 $(x_n, x_n + x)$ (단, x 는 임무시간(Mission time))동안 소프트웨어의 고장이 일어나지 않을 확률인 신뢰도 $\hat{R}(x | x_n)$ 는 다음과 같이 됨이 알려져 있다[6, 10].

$$\begin{aligned} \hat{R}(x | x_n) &= \exp \left(- \int_{x_n}^{x_n+x} \lambda(\tau) d\tau \right) \\ &= \exp \left[- \{ m(x+x_n) - m(x_n) \} \right] \end{aligned} \quad (27)$$

따라서 로그 포아송 실행시간모형에 대한 신뢰도는 평균값함수 (7)식과 $t = x_n$ 을 이용하면 다음과 같이 표현된다[5, 11].

$$\begin{aligned} R(x | T_R) &= \exp \left(- \frac{1}{\theta} [\ln(\lambda_0 \theta (x+t) + 1) - \ln(\lambda_0 \theta t + 1)] \right) \end{aligned} \quad (28)$$

따라서 소프트웨어 방출시간 T_R 이 신뢰도 $R_0 = \hat{R}(x | T_R)$ 을 확보해야 한다면 다음 방정식을 만족해야 한다.

$$\ln R_0 = - \frac{1}{\theta} [\ln(\lambda_0 \theta (x + T_R) + 1) - \ln(\lambda_0 \theta T_R + 1)] \quad (29)$$

한편, 소프트웨어 방출시간을 T 로 표현하고 $m(T)$ 와 $m(\infty)$ 을 각각 $(0, T]$ 와 $(0, \infty)$ 의 기간에 발견된 기대 고장수라고 표현하고 $C(T)$ 을 소프트웨어 라이프사이클(life cycle) 동안에 기대되는 소프트웨어 비용이라고 하면 $C(T)$ 는 다음과 같이 표현된다.

$$C(T) = c_1 m(T) + c_2 [m(\infty) - m(T)] + c_3 T \quad (30)$$

위 식에서 c_1 는 테스트 동안에 하나의 고장을 수리하는 비용이고 c_2 가동 중에 하나의 고장을 수리하는 비용($c_2 > c_1$), 그리고 c_3 는 단위 시간당 테스트 비용을 나타낸다. 이와 관련하여 총비용의 최소화는 무한고장 평균값 함수를 가진 NHPP 모형에 대하여 발생할 수 있다. 무한 수명에 대한 비용함수 $C(T)$ 식인 (27)식에서 $m(\infty)$ 은 직접 추정할 수 없기 때문에 이 식을 사용하기 위해서는 소프트웨어 수명시간인 T_{LC} 을 지정하여 분석한다[12]. 이러한 T_{LC} 는 소프트웨어마다 서로 다른 임의의 값이기 때문에 유한 고장 NHPP 모형이라고 할 수는 없다.

따라서 비용함수를 고려하여 소프트웨어의 모든 수명에서 총비용을 최소화함으로서 최적 테스트 시간을 결정할 수 있고 다음과 같은 식을 만족하면 비용함수 $C(T)$ 는 유일한 최소값을 가진다.

$$\frac{dC(T)}{dT} = 0, \frac{d^2 C(T)}{dT^2} > 0 \quad (31)$$

결국 소프트웨어 지정된 수명 T_{LC} 을 이용한 로그 포아송 실행시간모형 비용함수 $C(T)$ 는 다음과 같이 유도된다.

$$\begin{aligned} C(T) &= c_1 m(T) + c_2 [m(\infty) - m(T)] + c_3 T \quad (32) \\ &= (c_1 - c_2) \frac{1}{\theta} \ln(\lambda_0 \theta T + 1) \\ &\quad + \frac{c_2}{\theta} \ln(\lambda_0 \theta T_{LC} + 1) + c_3 T \end{aligned}$$

T 에 관해서 비용함수 $C(T)$ 을 미분하면 다음과 같은 방정식을 만족하는 최적방출시간 T_C 을 계산 할 수 있다.

$$\frac{(c_1 - c_2) \lambda_0}{\lambda_0 \theta T_C + 1} + c_3 = 0 \quad (33)$$

위 식에서 최적방출시간은 소프트웨어 지정 수명시간인 T_{LC} 와 의존하지 않는다는 것을 알 수 있다. 이러한 사실은 무한 고장 평균값 함수를 가진 NHPP모형들이 새로운 결점들이 발생함으로써 몇 개의 고장이 야기 될 수 있는 점을 고려한 모형으로 적합 시킬 수 있다[4, 12]. 실제로 만족할 만한 신뢰도가 부여되고 동시에 시스템 고장과 연계된 기대 총비용을 최소화시키기 위하여 필요하다면 충분한 테스트를 계속해야 한다. 따라서 신뢰성 요구를 만족하고 총 비용을 최소화하는 상황이 최적 방출 시간이다. 따라서 로그 포아송 실행시간모형을 사용한 최적 방출시간 T_{OP} 는 T_R 과 T_C 에 대하여 다음을 만족한다.

$$T_{OP} = \text{Max}(T_C, T_R) \quad (34)$$

(36) 식에서 T_R 과 T_C 는 다음 두 방정식에 의해서 계산 될 수 있다[3].

$$\ln R_0 = -\frac{1}{\theta} [\ln(\lambda_0 \theta (x + T_R) + 1) - \ln(\lambda_0 \theta T_R + 1)] \quad (35)$$

$$\frac{(c_1 - c_2) \lambda_0}{\lambda_0 \theta T_C + 1} + c_3 = 0 \quad (36)$$

유사한 방법으로 로그 파우어 모형에서도 T_R 과 T_C 는 다음 두 방정식에 의해서 계산 될 수 있다[12].

$$\ln R_0 = -a(\ln^b(1 + (x + T_R)) - \ln^b(1 + T_R)) \quad (37)$$

$$(c_1 - c_2) \frac{a b \ln^{b-1}(1 + T_C)}{1 + T_C} + c_3 = 0 \quad (38)$$

IV. 제안된 어랑 강도함수모형

소프트웨어의 수명분포에서 형상모수는 수명분포의 비대칭성(왜도, skewness), 첨도(kurtosis)등 분포의 형상을 결정하는 모수로 중요한 위치를 차지하고 있다[8]. 본 논문에서는 감마분포에서 형상모수가 고정된 어랑 분포의 경우를 제안하고자 한다.

즉, 형상모수가 2를 가지는 Erlang 분포 모형에 대한 신뢰도는 (20)식과 (27)식에서 $t = x_n$ 라고 하면 다음과 같이 표현된다.

$$\begin{aligned} R(x | T_R) & \quad (39) \\ &= \exp[-\beta x + \ln(1 + \beta(T_R + x)) - \ln(1 + \beta T_R)] \end{aligned}$$

따라서 소프트웨어 방출시간 T_R 이 신뢰도

$R_0 = \hat{R}(x | T_R)$ 을 확보해야 한다면 다음 방정식을 만족해야 한다.

$$\ln R_0 = -\beta x + \ln(1 + \beta(T_R + x)) - \ln(1 + \beta T_R) \quad (40)$$

단, 신뢰도 $R_0 = \hat{R}(x | T_R)$.

최소비용 관련 방출시간은 (20)식과 (30)를 연관하면 다음과 같이 유도 된다.

$$(c_1 - c_2) \left(\beta - \frac{\beta}{1 + \beta T_C} \right) + c_3 = 0 \quad (41)$$

따라서 Erlang 분포 모형을 사용한 최적 방출시간 T_{OP} 는 T_R 과 T_C 에 대하여 다음을 만족한다.

$$T_{OP} = \text{Max}(T_C, T_R) \quad (42)$$

유사한 방법으로 형상모수가 3인 분포 모형에 대한 신뢰도는 (29)식에서 $t = x_n$ 라고 하면 다음과 같이 표현된다.

$$\begin{aligned} \ln R_0 = & -\beta x + \ln \left((1 + \beta(T_R + x)) + \frac{\beta^2}{2}(x + T_R) \cdot (x + T_R) \right) \\ & - \ln \left(1 + \beta T_R + \frac{\beta^2}{2} \cdot T_R^2 \right) \end{aligned} \quad (43)$$

단, 신뢰도 $R_0 = \hat{R}(x | T_R)$.

최소비용 관련 방출시간도 $\alpha = 2$ 인 경우와 유사하게 다음과 같이 유도 된다.

$$(c_1 - c_2) \left(\beta - \frac{\beta + \beta^2 T_C}{1 + \beta T_C + \frac{\beta^2 T_C^2}{2}} \right) + c_3 = 0 \quad (44)$$

V. 수치적인 예

NTDS(Naval Tactical Data System)에 의해 발생된 소프트웨어 고장간격자료[9]를 적용한 감마형 분포 특수모형인 Erlang 분포 모형에 근거한 최적 방출시기를 분석하고자 한다.

이 자료는 250 시간단위에 26번의 고장이 발생된 자료이며 표 1에 나열 되어 있고 표2에서는 이 자료에 대한 기초통계량이 요약되어 있다. 이 표에서 왜도와 첨도가 각각 4.21과 19.07로 나타나 비대칭 분포임을 보여주고 있다. 그리고 제시하는 신뢰 모형들을 분석하기 위하여 우선 자료에 대한 추세 검정이 선행 되어야 한다.

추세 분석에는 일반적으로 라플라스 추세 검정(Laplace trend test)등이 사용된다. 이 검정을 실시한 결과 그림 2에서 라플라스 요인(Factor)이 -2와 2사이에 존재함으로써 신뢰성장(Reliability growth) 속성을 나타내고 있다. 따라서 이 자료를 이용하여 신뢰도와 소프트웨어 방출시기를 추정하는 것이 가능하다.

표 2에서는 이 자료에 대한 기초통계량을 요약하였고 그림 3에서의 감마분포의 확률도 그림도 거의 유사한 누적분포를 나타내고 있다. 모든 모형에 대한 모수 추정은 최우추정법을 이용하였고 비선형 방정식의 계산방법은 수치해석적 기본 방법인 이분법(Bisection method)을 사용하였다.

이러한 계산은 초기값을 10^{-4} 와 50을, 허용 한계(Tolerance for width of interval)는 10^{-10} 을 주고 수렴성을 확인 하면서 충분한 반복 횟수인 100번을 C-언어를 이용하여 모수 추정을 수행하였다. 따라서 모수 추정의 결과는 표 3에 요약되었다. 표 3에서는 $c_1 = 2(\$)$, $c_2 = 15(\$)$ 그리고 $c_3 = 0.5(\$)$ 라고 가정하고 시스템 수명시간은 2000시간이고 임무시간 x 는 1.5이고 R_0 을 0.95(95%)를 투입하여 각 모형에 대한 추정시간의 결과와 최적방출시간은 표 3에 요약되었다.

이 표에서 제안한 어랑 강도 모형($\alpha = 2$ 와 $\alpha = 3$ 인 경우)은 기존의 로그 포아송 모형과 로그 파우어 모형 보다 최적방출시간이 길지 않기 때문에 효율적이므로 상대적으로 우수한 모형으로 나타나고 있다.

표 1. 고장 간격자료(단위: 시간)

Table 1. Date of failure interval time(unit: hours)

Failure number	Failure Interval Time	Failure Time
1	9	9
2	12	21
3	11	32
4	4	36
5	7	43
6	2	45
7	5	50
8	8	58
9	5	63
10	7	70
11	1	71
12	6	77
13	1	78
14	9	87
15	4	91
16	1	92
17	3	95
18	3	98
19	6	104
20	1	105
21	11	116
22	33	149
23	7	156
24	91	247
25	2	249
26	1	250

표 2. 고장 간격자료의 기초통계량

Table 2. Basic statistics of failure interval data

평균	표준편차	최소값	Q1	중위수
9.62	17.79	1	2	5.5
Q3	최대값	왜도	첨도	관측 수
9	91	4.21	19.07	26

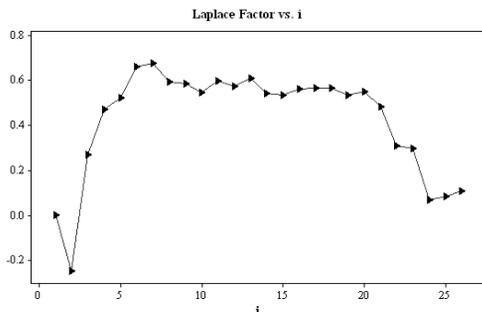


그림 2. 라플라스 추세검정
Fig. 2. Laplace trend test

로그 포아송 모형은 비용 측면 로그파우어 모형은 신뢰도 측면이 늦게 나타나고 있고 모두 유사한 방출 시간을 나타내고 있지만 형상모수가 $\alpha = 2.0$ 와 $\alpha = 3.0$ 인 경우의 어랑 모형은 신뢰도 측면에서의 방출시간은 빠르게 나타났지만 비용측면에서는 보다 늦게 나타나 이모형을 사용하고자 하면 비용 측면에 대한 관리가 절실히 필요한 모형으로 사료된다.

표 2. 각 모형의 모수 추정값

Table 2. Parameter estimation of each model

Model	MLE	
Log Poisson execution time	$\hat{\theta} = 0.000312$	$\hat{\lambda}_0 = 26.02236$
Log Power	$\hat{a} = 34.52284$	$\hat{b} = 0.528786$
Erlang distribution	$\alpha = 2.0$	$\hat{\beta} = 0.063948$
	$\alpha = 3.0$	$\hat{\beta} = 0.053232$

표 3. 최적 방출시간 추정($R_0 = 0.95$)

Table 3. Estimation of optimal release time($R_0 = 0.95$)

Model	추정시간	최적방출시간(T_{OP})
Log Poisson execution time	$\hat{T}_R = 31922.8321$ $\hat{T}_C = 83209.9224$	83209.9224
Log Power	$\hat{T}_R = 62982.9573$ $\hat{T}_C = 64.9398$	62982.9573
Erlang distribution	$\alpha = 2.0$ $\hat{T}_R = 2.0545$ $\hat{T}_C = 23.5988$	23.5988
	$\alpha = 3.0$ $\hat{T}_R = 6.9861$ $\hat{T}_C = 114.2467$	114.2467

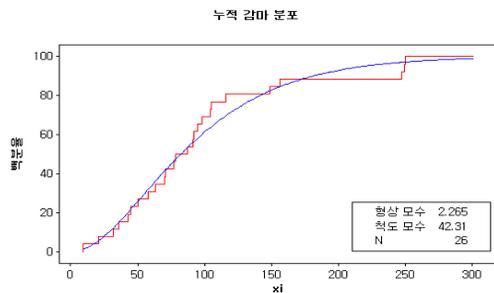


그림 3. 감마분포를 적용한 확률도
Fig. 3. Probability based on gamma distribution

결국 제안된 어랑 강도 모형은 기존의 로그 포아송 모형과 로그 파우어 모형 비해 상대적으로 최적방출 시간이 효율적임을 보여주고 있지만 이 결과는 적합 시킨 자료가 그림3에 보여 주듯이 감마분포(형상모수가 정수인 경우 어랑분포)에 따르는 결과이기 때문에 다른 자료를 적합 시키면 결과는 달라질 수 있다. 따라서 감마형 분포의 특수모형인 어랑 모형도 이 분야에 새로운 모형으로 선택 할 수 있음을 보여주고 있다.

V. 결론

본 연구는 형상모수를 고정시킨 어랑 모형을 적용한 무한고장 NHPP 모형을 이용하여 최적 방출시기에 관한 문제를 알아보았다. 즉, 대용량 소프트웨어가 수 경과 변경하는 과정에서 결점의 발생을 거의 피할 수 없는 상황이 현실이다. 실제로 만족할 만한 신뢰도가 부여되고 동시에 시스템 고장과 연계된 기대 총비용을 최소화시키기 위하여 필요하다면 충분한 테스트를 계속해야 한다.

따라서 신뢰성 요구를 만족하고 총 비용을 최소화하는 상황이 최적 방출 시간이다. 본 연구에서는 어랑 강도모형을 적용한 방출시기 모형을 시도한 결과 효율성을 입증 하였다. 따라서 어랑 분포모형도 이 분야에서 적용 가능한 모형이 될 수 있음을 확인하였다.

경우에 따라서는 왜도와 첨도 측면에서 효율적인 카파분포, 지수화지수분포 등 업데이트된 분포에 대한 방출 시기 문제를 비교 분석하는 연구도 가치 있는 일이라 판단되고 이 연구를 통하여 소프트웨어 개발자들은 방출최적시기를 파악 하는데 어느 정도 도움을 줄 수 있으리라 사료 된다.

참고문헌

- [1] 김희철, “극값 분포 특성을 가진 소프트웨어 신뢰성 보증 모형에 관한 연구”, *한국 통신학회 논문지*, 34 권6호, pp. 623-629, 2009.
- [2] 김희철, “지수화 지수 분포에 의존한 NHPP 소프트웨어 신뢰성장 모형에 관한 연구”, *한국 컴퓨터 정보학회 논문지*, 제11권 5호, pp. 9-18, 2006.
- [3] 김희철 “어랑분포에 근거한 소프트웨어 최적방출시기에 관한 연구”, *남서울대학교 논문집*, 제14권, pp. 59-74, 2008.
- [4] 김희철, 박형근 “와이블분포특성에 근거한 소프트웨어 최적방출시기에 관한 비교 연구”, *한국 산학기술학회 논문지*, 제10권8호, pp. 1903-1910, 2009.
- [5] Almering, V. and Genuchten, M, V and Cloudt, G. and Sonnemans, P. J. M, “Using Software Reliability Growth Models in Practice”, *IEEE SOFTWARE*, pp. 82-88, 2007.
- [6] Gokhale, S. S. and Trivedi, K. S. “A time/structure based software reliability model”, *Annals of Software Engineering*, 8, pp. 85-121. 1999.
- [7] Huang, C. Y, “Cost-Reliability-optimal release policy for software reliability models incorporating improvements in testing efficiency”, *The journal of Systems and software*, Vol, 77, pp. 139-155, 2005.
- [8] K. Kanoun and J. C. Laprie, “Handbook of Software Reliability Engineering”, M.R. Lyu, Editor, chapter trend Analysis. McGraw-Hill New York, NY, pp.401-437, 1996.
- [9] Kuo, L. and Yang, T. Y, “Bayesian Computation of Software Reliability”, *Journal of the American Statistical Association*, Vol.91, pp.763-773, 1996.
- [10] Lawless, J. F, “Statistical Models and Methods for Lifetime Data”, John Wiley & Sons, New York, 1981.
- [11] Musa, J. D and Okumoto, K, “A Logarithmic Poisson Execution Time Model for Software Reliability Measurement”, *Proceeding the 7th International Conference on Software engineering*, pp. 230-238, 1984.
- [12] Xie, M. and Homg, G. Y, “Software release time determination based on unbound NHPP model”,

Proceeding of the 24th International Conference on Computers and Industrial Engineering. pp. 165-168, 1999.

- [13] Yang, B. and Xie, M, "A study of operational and testing reliability in software reliability analysis", RELIABILITY ENGINEERING & SYSTEM SAFETY, Vol, 70, pp,323-329, 2000.
- [14] 이레테크, "새 MINITAB 실무완성", 이레테크 미니탭 사업부, 2005.
- [15] 임희균, "온톨로지 기반의 로봇 컴포넌트 동적 재구성에 관한 연구", 한국지식정보기술학회 논문지, 제4권 제4호, pp.1-8, 2009.



김희철(Hee-Cheul Kim)

2002년 동국대학교 대학원 통계학과
(이학석사)
2008년 동국대학교 대학원 통계학과
(이학박사)

2005년~현재 남서울대학교 산업경영공학과 교수

※ 관심분야: 소프트웨어신뢰성공학, 전산통계