

# **초지수분포(Hyper-exponential)를 이용한 소프트웨어 신뢰성장 모형에 관한 연구**

김희철\* · 신현철\*\*

## **요 약**

유한고장수를 가진 비동질적인 포아송 과정에 기초한 모형들에서 잔존 결합 1개당 고장 발생률은 일반적으로 상수, 혹은 단조증가 및 단조 감소 추세를 가지고 있다. 본 논문에서는 기존의 소프트웨어 신뢰성 모형인 Goel-Okumoto 모형과 Yamada-Ohba-Osaki 모형을 재조명하고 이 분야에 적용될수 있는 hyper-exponential 분포를 이용한 모형을 제안하였다. 수치적인 예에서는 Minitab(version 14) 통계 페키지에 있는 와이블분포(형상모수가 0.5이고 척도모수가 1)에서 발생시킨 30개의 난수를 이용한 모의 실험 고장 간격시간으로 구성된 자료를 이용하였고 모수추정 방법은 최우추정법과 일반적인 수치해석 방법인 이분법을 사용하여 모수 추정을 실시하였다. 그리고 모형 설정과 선택 판단기준은 편차 자승합을 이용한 적합도 검정이 사용되었다.

## **The Study for NHPP Software Reliability Growth Model Based on Hyper-exponential Distribution**

Hee-Cheul Kim\* · Hyun-Cheul Shin\*\*

## **ABSTRACT**

Finite failure NHPP models presented in the literature exhibit either constant, monotonic increasing or monotonic decreasing failure occurrence rates per fault. In this paper, Goel-Okumoto and Yamada-Ohba-Osaki model was reviewed, proposes the hyper-exponential distribution reliability model, which maked out efficiency application for software reliability. Algorithm to estimate the parameters used to maximum likelihood estimator and bisection method. For model determination and selection, explored goodness of fit (the error sum of squares). The methodology developed in this paper is exemplified with a software reliability random data set introduced by of Weibull distribution (shape 0.1 & scale 1) of Minitab (version 14) statistical package.

**Key words :** Software Reliability Model, NHPP, Hyper-exponential Distribution, SSE

\* 남서울대학교 산업경영공학과

\*\* 백석문화대학 컴퓨터정보학부

## 1. 서 론

소프트웨어 테스팅 단계에서 소프트웨어 고장 수(Number of failure)와 고장간격시간에 의해 소프트웨어 고장현상을 수리적으로 모형화 하면 소프트웨어에 대한 평가를 보다 쉽게 할 수 있으며 신뢰도 모형에 의해 소프트웨어 고장수, 소프트웨어 고장발생간격시간, 소프트웨어 신뢰도 및 고장을 등의 신뢰성 평가측도들이 추정되어 미래의 고장시간을 예측할 수도 있다.

소프트웨어 고장시간은 수명자료가 된다. 따라서 비음(Nonnegative)의 값을 가지기 때문에 이 분야에서는 주로 지수분포, 와이블분포, 감마분포 등 일반화 감마 분포(혹은 어랑분포)가 많이 사용되어 왔다[1]. 그러나 이러한 감마분포 외에도 통계자료 해석시 자료를 대수변환 이 후에 정규분포로 처리하면 되는 대수정규분포(Lognormal distribution), 이 분포와 유사한 대수로지스틱 분포(Log-logistic distribution)도 이 분야에 사용이 가능하다[3]. 또, 자유도(Degree of freedom ; df)에 의존하는 카이제곱( $\chi^2$  distribution) 분포도 적용이 가능하다. 이러한 카이제곱 분포는 변화가 일어날 때 까지의 대기시간을 나타내는 데 많이 사용되는 감마분포(Gamma Distribution)의 특수한 경우로도 접근 할 수 있기 때문에 소프트웨어 신뢰성 수명분포로 설명 할 수 있다[2].

본 논문에서는 분포함수가 효율적으로 사용할 수 있는 Hyper-exponential 분포[10]를 적용하고자 한다. 이러한 Hyper-exponential 분포를 수명분포로 사용한 NHPP 모형에 대한 신뢰성 척도를 추정하고 이를 바탕으로 형상모수에 따른 모형의 효율성과 그 특성을 알아보고자한다.

본 논문의 제 2장에서는 이론적 배경으로서 유한 고장 NHPP 모형에 대하여 서술하였고 제 3장에서는 Hyper-exponential 분포 신뢰성 모형 및 최우추정법을 이용한 모수추정에 대하여 설명하고 제 4장에서는 모의 고장자료를 이용하여 각 모형

에 대한 모수추정 및 모형비교를 실시하였으며 마지막으로 제 5장에서는 결론을 나열 하였다.

## 2. 이론적 배경에 대한 고찰

신뢰도에서 관측시간  $(0, t]$  사이에 발견된 고장 수  $N(t)$ 를 모형화 하는데 비동질적 포아송 과정(Non-homogeneous poission process ; NHPP)이 널리 사용하여 왔다. 이 과정(Process)에서 강도함수(Intensity function) 혹은 고장 발생률(Rate of occurrence of failure ; ROCOF)  $\lambda(t) = dE[N(t)]/dt$  은  $t$ 에 대한 단조(Monotonic)함수로 흔히 가정 한다 [2]. 이 범주에서 지금까지 알려진 모형들은 Goel-Okumoto 모형, Weibull 모형 그리고 Cox-Lewis 모형등이 있는데 이 모형들에 대한 강도함수는 각각 시간에 의존한 함수, 힘(Power) 함수, 대수선형(Log-linear) 함수를 가정하였다[6].

NHPP 모형에서 평균값 함수  $m(t)$  (Mean value function)와 강도 함수  $\lambda(t)$ 는 다음과 같은 관계로 표현 할 수 있다[4].

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

$N(t)$ 는 모수  $m(t)$ 를 가진 포아송 확률밀도함수(Probability density function)로 알려져 있다. 즉,

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

이처럼 시간관련 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 나타난다. 이러한 NHPP 모형들은 유한 고장 모형과 무한 고장 범주로 분류한다[8]. 유한 고장(Finite failure) NHPP 모형들은 충분한 테스트 시간이 주어지면 결함들(Faults)의 기대값이 유한

값  $(\lim_{t \rightarrow \infty} m(t) = \theta < \infty)$  을 가지고 반면에 무한 고장 (Infinite failure) NHPP 모형들은 무한 값을 가진다고 가정 된다. 유한 고장 NHPP 모형에서 충분한 테스트 시간이 주어졌을 때 탐색되어 질 수 있는 결합의 기대값을  $\theta$ 라고 표현하고  $F(t)$ 를 분포 함수라고 표현하면 유한 고장 NHPP 모형의 평균 값 함수는 다음과 같이 표현 할 수 있다[3].

$$m(t) = \theta F(t) \quad (3)$$

식 (2), 식 (3)으로 부터 순간고장 강도함수(Instantaneous failure intensity)  $\lambda(t)$ 는 다음과 같이 유도된다.

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

식 (4)를 다음과 같이 변형하여 표기 할 수도 있다.

$$\lambda(t) = [\theta - m(t)] \frac{F'(t)}{1 - F(t)} = [\theta - m(t)] h(t) \quad (5)$$

단,  $h(t) = \frac{F'(t)}{1 - F(t)}$  는 위험함수(Hazard function, 고장률 함수)으로서 소프트웨어 결합당 고장 발생률을 의미하고  $[\theta - m(t)]$ 은  $t$  시점에서 소프트웨어에 남아있는 결합들의 기대값을 나타낸다.  $[\theta - m(t)]$ 의 값은 시점  $t$ 에 대한 단조 비증가 함수(Monotonically nonincreasing function)가 된다. 즉, 시간이 지남에 따라 그 고장을 찾아 제거하는 디버깅 과정을 거치면서 제거되기 때문에 감소성을 가진다.

$\lambda(t)$ 는  $h(t)$ 의 속성에 따라 달라지며 그 추세는 상수나 증가, 혹은 감소(증가)하다가 증가(감소)하는 패턴을 가질 수도 있다[2].

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

$$x_k = \sum_{i=1}^k t_i \quad (k=1, 2, \dots, n; 0 \leq x_1 \leq x_2 \leq \dots \leq x_n) \quad (6)$$

이라고 하면 데이터 집합  $D_t$ 는  $\{n, x_1, x_2 \dots x_n ; t\}$

와 같이 구성된다.  $n$ 번째까지 고장시점이 관찰된 고장 절단 모형일 경우에 데이터 집합  $D_{x_n}$ 은  $\{x_1, x_2 \dots x_n\}$ 으로 구성되며 이 시간 절단 모형에서의  $\theta$ 를 모수공간이라고 표시하면 우도함수는 다음과 같이 알려져 있다[4, 5].

$$L_{NHPP}(\Theta | D_{x_n}) = \left( \prod_{i=1}^n \lambda(x_i) \right) \exp(-m(x_n)) \quad (7)$$

이 분야의 기본적 모형인 Goel-Okumoto 모형 [9]은  $h(t)$ 가 정수 패턴을 가짐으로서 시점  $t$ 에 독립이고 잘 알려진 Yamada, Ohba-Osaki 모형은 단조 비감소 패턴을 가진다[2, 3].

NHPP 모형에서 테스트 시점  $x_n$ 에서 소프트웨어 고장이 일어난다고 하는 가정하에서 신뢰구간  $(x_n, x_n + t]$  (단,  $t$ 는 임무시간(Mission time)) 사이에서 소프트웨어의 고장이 일어나지 않을 확률인 신뢰도(Reliability)  $\hat{R}(t|x_n)$ 는 다음과 같이 됨이 알려져 있다[3, 5].

$$\begin{aligned} \hat{R}(t | x_n) &= e^{- \int_{x_n}^{x_n+t} \lambda(\tau) d\tau} \\ &= \exp[-\{m(t+x_n) - m(x_n)\}] \end{aligned} \quad (8)$$

## 2.1 Hyper-exponential 분포

이 장에서 hyper-exponential 분포[10]에 대하여 요약 하고자 한다.

확률통계부분에서 hyper-exponential 분포는 독립적인 지수분포의 가중합으로 정의된다 이 분포의 확률밀도 함수와 분포함수는 다음과 같이 표현됨이 알려져 있다.

$$f_X(x) = \sum_{i=1}^n f_{X_i}(y) p_i \quad (9)$$

단,  $X_i$ 는 독립적으로 모수  $\omega_i$  ( $i=1, 2, \dots, n$ )를 가지는 지수분포를 따르고  $n$ 은 양의 상수이며  $\sum_{i=1}^n p_i = 1$ 을 만족한다.

$$F_X(x) = \sum_{i=1}^n F_{X_i}(x_i) p_i \quad (10)$$

$$\text{단, } f(x_i) = \omega_i e^{-\omega_i x_i} (x_i \geq 0),$$

$$F(x_i) = 1 - e^{-\omega_i x_i} (x_i \geq 0)$$

이러한 혼합 분포 모형은 부품이나 시스템 집단이 수명분포가 서로 다른  $n$ 개의 부분집단으로 구성되는 경우가 이에 해당한다.

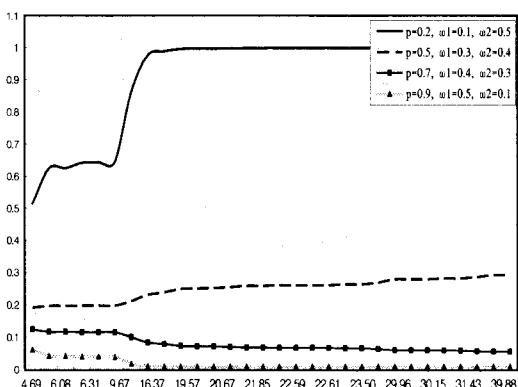
위험함수는 식 (9)와 식 (10)과 관련하면 다음과 같이 유도 된다.

$$h(t | p_i, \omega_i) = \frac{\sum_{i=1}^n p_i f(t | \omega_i)}{1 - \sum_{i=1}^n p_i F(t | \omega_i)} \quad (11)$$

예를 들어  $n=2$ 인 경우에 위험함수의 형태는 다음과 같은 형태가 된다.

$$h(t | p, \omega_1, \omega_2) = \frac{p(\omega_1 e^{-\omega_1 t} - \omega_2 e^{-\omega_2 t}) + \omega_2 e^{-\omega_2 t}}{p(e^{-\omega_1 t} - e^{-\omega_2 t}) + e^{-\omega_2 t}}$$

(그림 1)에서 보여 주듯이  $p$ 가 0.5보다 작은 경우는 단조증가형태이고  $p$ 가 0.5보다 큰 경우는 단조감소하는 형태가 된다.



(그림 1) hyper-exponential 분포의 위험함수

## 2.2 기존의 신뢰성 모형

기존의 모형은 주로 수명 분포가 다음과 같은 감마분포의 분포함수를 사용하고 있다[4].

$$F_{G_a}(t | k, \beta) = 1 - e^{-\beta t} \left( \sum_{i=0}^{k-1} \frac{(\beta t)^i}{i!} \right) \quad (12)$$

$(t, k > 0, \beta > 0)$

식 (12)에서 감마분포의 확률밀도함수는 다음과 같이 유도 할 수 있다[2].

$$f_{G_a}(t | k, \beta) = \frac{\beta^k}{\Gamma(k)} t^{k-1} e^{-\beta t} \quad (k, \beta > 0, t > 0) \quad (13)$$

따라서 식 (12)와 식 (13)에 관련하여 소프트웨어 결함(Fault)당 고장(Failure) 발생률이 일정하거나 증가 혹은 감소하는 특징을 가지는 모형이 감마 모형이 되고 유한 고장 NHPP 모형의 평균값 함수  $m(t)$ 와 강도함수  $\lambda(t)$ 는 다음과 같이 표현 할 수 있다.

$$m(t) = \theta F_{G_a}(t | k, \beta) = \theta \left[ 1 - e^{-\beta t} \left( \sum_{i=0}^{k-1} \frac{(\beta t)^i}{i!} \right) \right] \quad (14)$$

$$\lambda(t) = \theta f_{G_a}(t | k, \beta) = \theta \frac{\beta^k}{\Gamma(k)} t^{k-1} e^{-\beta t} \quad (15)$$

식 (14)의 분포함수가 폐쇄형(closed form)의 형태이므로 기존의 모형은 주로 함수는  $k$  값이 양의 정수일 때만을 고려하고 있다.

이 분야에서 가장 기본적인 모형은 Goel-Okumoto 모형[6]이다. 이 모형은 결함당 고장발생 시간의 분포(수명 분포)를 지수분포를 가정하였다. 따라서 소프트웨어 결함(Fault)당 고장(Failure) 발생률이 일정한 형태를 가지며 식 (14)와 식 (15)에서 형상모수  $k=1$ 인 경우와 같이 된다. 이 모형을 개선한 모형이 잘 알려진 S 모형(Yamada, Ohba-Osaki 모형[7])이다. 따라서 이 모형은 수명 분포를 감마분포중에서 형상모수  $k=2$ 인 분포를 말한다.

### 2.2.1 Goel-Okumoto 모형과 Yamada-Ohba-Osaki 모형

Goel-Okumoto 모형[9]에 대한 평균값 함수는  $m(t|\theta, \beta_1) = \theta(1 - e^{-\beta_1 t})$  ( $\theta > 0, \beta_1 > 0$ )이라고 알려져 있다. 시간  $(0, t]$  까지 조사하기 위한 시간 절단(Time truncated)모형은  $n$  번째 까지 고장시점 자료를

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

이라고 하면 데이터 집합  $D_t$  는  $\{n, x_1, x_2, \dots, x_n; t\}$  와 같이 구성된다.

$t$ 을 최종 고장시점  $x_n$ 으로 대치하고 고장 발생률  $\lambda(t|\theta, \beta_1) = \theta \beta_1 e^{-\beta_1 t}$  을 이용하면 우도함수는 다음과 같이 표현 할 수 있다.

$$L_{GO}(\theta, \beta_1 | D_{x_n}) = \left( \prod_{k=1}^n \theta \beta_1 e^{-\beta_1 x_k} \right) \exp[-\theta(1 - e^{-\beta_1 x_n})] \quad (16)$$

최우추정법(MLE)을 이용하기 위한 로그우도함수를 구하면,

$$\ln L_{GO} = n \ln \theta + n \ln \beta_1 - \beta_1 \sum_{k=1}^n x_k - \theta(1 - e^{-\beta_1 x_n}) \quad (17)$$

으로 표현된다.

따라서 고장절단모형에서의 모수  $\theta$ 와  $\beta_1$ 에 관한 편미분식은 다음과 같이 유도할 수 있다.

$$\frac{\partial \ln L_{GO}}{\partial \theta} = \frac{n}{\theta} - 1 + e^{-\beta_1 x_n} = 0,$$

$$\frac{\partial \ln L_{GO}}{\partial \beta_1} = \frac{n}{\beta_1} - \sum_{k=1}^n x_k - \theta x_n e^{-\beta_1 x_n} = 0$$

각 모수에 대한 최우추정량  $\hat{\theta}_{MLE}$  와  $\hat{\beta}_1 MLE$  은 다음식을 만족한다.

$$\frac{n}{\hat{\theta}} = 1 - \exp(-\hat{\beta}_1 x_n), \quad (18)$$

$$\frac{n}{\hat{\beta}_1} = \sum_{k=1}^n x_k + \hat{\theta} x_n \exp(-\hat{\beta}_1 x_n) \quad (19)$$

식 (18)과 식 (19)를 수치적으로 풀어 두 모수를 구할 수 있다.

Yamada-Ohba-Osaki 모형[3, 7]에 대한 평균값 함수는  $m(t|\theta, \beta_2) = \theta [1 - (1 + \beta_2 t) e^{-\beta_2 t}]$  ( $\theta > 0, \beta_2 > 0$ )이라고 알려져 있고  $\lambda(t|\theta, \beta_2) = \theta \beta_2^2 t e^{-\beta_2 t}$  가 되고 우도함수는 다음과 같다.

$$L_{YOO}(\theta, \beta_2 | D_{x_n}) = \left( \prod_{k=1}^n \theta \beta_2^2 x_k e^{-\beta_2 x_k} \right) \exp[-\theta [1 - (1 + \beta_2 x_n) e^{-\beta_2 x_n}]] \quad (20)$$

따라서 로그우도함수를 구하면 다음과 같이 표현된다.

$$\begin{aligned} \ln L_{YOO} = & n \ln \theta + 2n \ln \beta_2 + \sum_{k=1}^n \ln x_k - \beta_2 \sum_{k=1}^n x_k \\ & - \theta [1 - (1 + \beta_2 x_n) e^{-\beta_2 x_n}] \end{aligned} \quad (21)$$

따라서 고장절단 모형에서의 모수  $\theta$ 와  $\beta_2$ 에 관한 편미분식은 다음과 같은 형태로 유도된다.

$$\frac{\partial \ln L_{YOO}}{\partial \theta} = \frac{n}{\theta} - 1 + e^{-\beta_2 x_n} + \beta_2 x_n e^{-\beta_2 x_n} = 0,$$

$$\frac{\partial \ln L_2}{\partial \beta_{YOO}} = \frac{2n}{\beta_2} - \sum_{k=1}^n x_k - \theta \beta_2 x_n^2 e^{-\beta_2 x_n} = 0$$

따라서 각 모수에 대한 최우추정량은  $\hat{\theta}_{MLE}$  와  $\hat{\beta}_{2 MLE}$  은 다음 식을 만족한다.

$$\frac{n}{\hat{\theta}} = 1 - \exp(-\hat{\beta}_2 x_n) - \hat{\beta}_2 x_n \exp(-\hat{\beta}_2 x_n) \quad (22)$$

$$\frac{2n}{\hat{\beta}_2} = \sum_{k=1}^n x_k + \hat{\theta} \hat{\beta}_2 x_n^2 \exp(-\hat{\beta}_2 x_n) \quad (23)$$

식 (22)와 식 (23)을 수치적으로 풀어 두 모수를 구할 수 있다.

### 3. 제안된 hyper-exponential 분포를 이용한 신뢰성 모형

hyper-exponential 분포모형에 대한 평균값 함수는 식 (3)과 식 (4)를 연관하여 평균값 함수와 고장 발생률  $\lambda_{Hyper}(t|p_i, \omega_i)$ 을 다음과 같이 유도 할 수 있다.

$$m_{Hyper}(t|\theta, p_i, \omega_i) = \theta \sum_{i=1}^n F_{X_i}(\omega_i) p_i \quad (24)$$

$$\lambda_{Hyper}(t|\theta, p_i, \omega_i) = \theta \sum_{i=1}^n f_{X_i}(\omega_i) p_i \quad (25)$$

$$\text{단, } t > 0, \theta > 0, \omega > 0, \sum_{i=1}^n p_i = 1.$$

$n=2$ 인 경우에는 다음과 같이 표현 할 수 있다.

$$m(t|\theta, p, \omega_1, \omega_2) = \theta - \theta p(e^{-\omega_1 t} - e^{-\omega_2 t}) - \theta e^{-\omega_2 t}$$

$$\lambda(t|\theta, p, \omega_1, \omega_2) = \theta p(\omega_1 e^{-\omega_1 t} - \omega_2 e^{-\omega_2 t}) + \theta \omega_2 e^{-\omega_2 t}$$

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

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

이라고 하면 데이터 집합  $D_t$ 는  $\{n, x_1, x_2, \dots, x_n; t\}$  와 같이 구성된다.

$t$ 을 최종 고장시점  $x_n$ 으로 대치하고 평균값 함수  $m(t|\theta, p, \omega_1, \omega_2)$ 와 고장발생률  $\lambda(t|\theta, p, \omega_1, \omega_2)$ 을 이용하면 우도함수  $L(\theta, p, \omega_1, \omega_2 | D_{x_n})$ 는 다음과 같이 표현 할 수 있다.

$$L(\theta, p, \omega_1, \omega_2 | D_{x_n})$$

$$= \prod_{i=1}^n \theta [p(\omega_1 e^{-\omega_1 x_i} - \omega_2 e^{-\omega_2 x_i}) + \omega_2 e^{-\omega_2 x_i}]$$

$$\cdot \exp[-\theta \{1 - p(e^{-\omega_1 x_n} - e^{-\omega_2 x_n}) + \omega_2 e^{-\omega_2 x_n}\}]$$

따라서 최우추정법을 이용하기 위한 로그우도 함수는 다음과 같이 유도된다.

$$\ln L(\theta, p, \omega_1, \omega_2 | D_{x_n}) =$$

$$n \ln \theta + \sum_{i=1}^n \ln [p(\omega_1 e^{-\omega_1 x_i} - \omega_2 e^{-\omega_2 x_i}) + \omega_2 e^{-\omega_2 x_i}]$$

$$\cdot [-\theta \{1 - p(e^{-\omega_1 x_n} - e^{-\omega_2 x_n}) + \omega_2 e^{-\omega_2 x_n}\}] \quad (26)$$

가중치  $p$  값은 상수(사전에 알고 있는 경우)라고 가정 했을 때 최우추정법을 이용하기 위하여 식 (26)을  $\theta$ 와  $\omega_1, \omega_2$ 에 대하여 편미분을 하면 다음과 같은 식을 유도 할 수 있다.

$$\frac{\partial \ln L(\theta, p, \omega_1, \omega_2 | D_{x_n})}{\partial \theta} =$$

$$\frac{n}{\theta} - \{1 - p(e^{-\omega_1 x_n} - e^{-\omega_2 x_n}) + \omega_2 e^{-\omega_2 x_n}\} = 0 \quad (27)$$

$$\frac{\partial \ln L(\theta, p, \omega_1, \omega_2 | D_{x_n})}{\partial \omega_1} =$$

$$\sum_{i=1}^n \frac{p(e^{-\omega_1 x_i} - \omega_1 x_i e^{-\omega_1 x_i})}{p(\omega_1 e^{-\omega_1 x_i} - \omega_2 e^{-\omega_2 x_i}) + \omega_2 e^{-\omega_2 x_i}} - x_n \cdot p \cdot e^{-\omega_1 x_n} = 0 \quad (28)$$

$$\frac{\partial \ln L(\theta, p, \omega_1, \omega_2 | D_{x_n})}{\partial \omega_2} =$$

$$\sum_{i=1}^n \frac{-p(e^{-\omega_2 x_i} - \omega_2 x_i e^{-\omega_2 x_i}) + e^{-\omega_2 x_i} - \omega_2 x_i e^{-\omega_2 x_i}}{p(\omega_1 e^{-\omega_1 x_i} - \omega_2 e^{-\omega_2 x_i}) + \omega_2 e^{-\omega_2 x_i}} - p \omega_2 e^{-\omega_2 x_n} + e^{-\omega_2 x_n} - \omega_2 x_i e^{-\omega_2 x_n} = 0 \quad (29)$$

따라서 고장절단 모형에서의 식 (27)과 식 (28), 식 (29)를 비선형 연립방정식(수치해석적 방법)을 이용하여 풀면 최우추정치  $\hat{\omega}_1^{MLE}$ 와  $\hat{\omega}_2^{MLE}$ ,  $\hat{\theta}^{MLE}$ 의 값을 구 할 수 있다.

본 논문에서는 Hyper-exponential 분포모형의 가중치  $p$  값과  $\omega_1$ 은 이분법의 수렴성을 위하여 미리 알고 있는 고정값을 이용하여 모형 모수값을 제시하고자 한다.

#### 4. 수치적인 예

이 장에서 Minitab(version 14) 통계 페키지에

있는 와이블분포(형상모수가 0.5이고 척도모수가 1)에서 발생시킨 30개의 난수를 이용한 모의 실험 고장 간격시간으로 구성된 자료를 이용하고자 한다. 이 자료는 <표 1>에 나열 되어 있다.

소프트웨어 신뢰성 모형의 모수 추정은 최우추정법을 이용하였고 비선형 방정식의 계산방법은 수치해석적 기본 방법인 이분법(Bisection method)을 사용하였다. 이러한 계산은 초기값을  $10^{-4}$ 와 10을, 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.

기존의 모형과 hyper-exponential 신뢰도 모형에 대한 모수의 추정값들의 결과는 <표 2>에 요약되었다. 이 표에서 hyper-exponential 모형에서 수치적인 수렴성 확보를 위하여 가중치  $p$ 와  $\omega_1$ 를 각각 0.1, 0.3, 0.5, 0.7, 0.9, 0.1, 0.2, 0.3, 0.4, 0.5를 할당하여 추정된 최우추정값들이다.

이 분야에서는 일반적으로 모형 선택의 하나의 방법으로 편차자승합(SSE [5, 8, 9])을 이용할 수 있다. 이러한 결과는 <표 3>에 요약되었다. 이 표에서 hyper-exponential 신뢰도 모형들이 가중치  $p = 0.7$ ,  $\omega_1 = 0.4$ ,  $p = 0.9$ ,  $\omega_1 = 0.5$ 에 따른 모형들이 이 분야에서 기준에 알려진 모형인 Yamada-Ohba-Osaki 모형이나 Goel-Okumoto 모형에 비해 상대적으로 효율적 모형으로 나타나고 있다.

평균값 함수에 대해 추정된 그림이 (그림 1)에서 보여주고 있다, 이 평균값 함수 패턴도  $p = 0.7$ ,  $\omega_1 = 0.4$ ,  $p = 0.9$ ,  $\omega_1 = 0.5$ 에 따른 Hyper-exponential 모형이 참값인 30에 근접하게 추정된 모습을 볼수 있고  $t$ 는 임무시간(Mission time))에 대한 즉, 소프트웨어의 고장이 일어나지 않을 확률인 신뢰도(Reliability)  $\hat{R}(t | x_n)$ 는 (그림 2)와 같다. 이 그림에서도  $p = 0.7$ ,  $\omega_1 = 0.4$ 에 따른 Hyper-exponential 모형이 상대적으로 효율적인 신뢰도를 나타내고 있다.

&lt;표 1&gt; 모의고장 자료

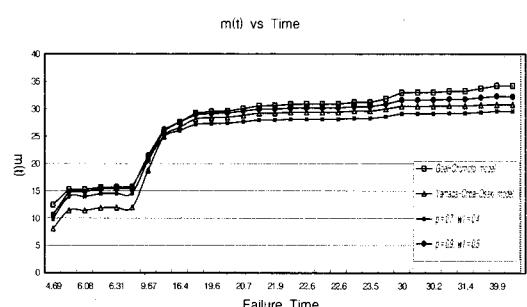
Failure number	Failure Time	Failure number	Failure Time
1	4.69	16	22.5
2	6.08	17	22.59
3	6.08	18	22.61
4	6.29	19	22.61
5	6.31	20	23.5
6	6.33	21	23.5
7	9.67	22	25.38
8	14.39	23	29.96
9	16.37	24	30.07
10	19.02	25	30.15
11	19.57	26	31.42
12	19.76	27	31.43
13	20.67	28	34.76
14	21.84	29	39.88
15	21.85	30	39.9

&lt;표 2&gt; 각 모형의 모수 추정값

Model			MLE
<i>Goel-Okumoto Model</i>			$\hat{\beta}_{1,MLE} = 0.0939$
			$\hat{\theta}_{MLE} = 35.025$
<i>Yamada-Ohba-Osaki Model</i>			$\hat{\beta}_{2,MLE} = 0.2124$
<i>Hyper-exponential Model</i>	$p = 0.1$		$\hat{\theta}_{MLE} = 30.8621$
	$\omega_1 = 0.1$		$\hat{\omega}_{2,MLE} = 0.3028$
			$\hat{\theta}_{MLE} = 29.726$
	$p = 0.3$		$\hat{\omega}_{2,MLE} = 0.1321$
	$\omega_1 = 0.2$		$\hat{\theta}_{MLE} = 34.726$
	$p = 0.5$		$\hat{\omega}_{2,MLE} = 0.2132$
	$\omega_1 = 0.3$		$\hat{\theta}_{MLE} = 30.726$
	$p = 0.7$		$\hat{\omega}_{2,MLE} = 0.1341$
	$\omega_1 = 0.4$		$\hat{\theta}_{MLE} = 29.8941$
	$p = 0.9$		$\hat{\omega}_{2,MLE} = 0.1132$
	$\omega_1 = 0.5$		$\hat{\theta}_{MLE} = 32.7264$

〈표 3〉 각 모형의 SSE 값

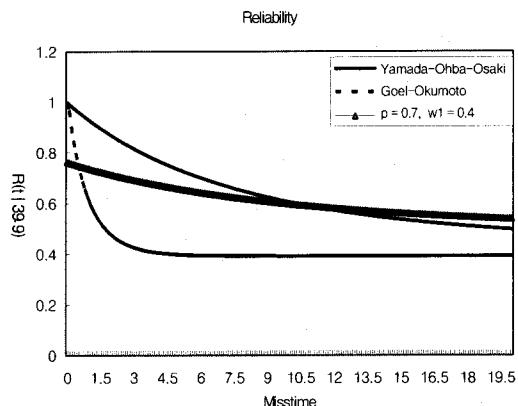
Model		SSE
<i>Goel-Okumoto Model</i>		4955.526
<i>Yamada-Ohba-Osaki Model</i>		3647.482
<i>Hyper-exponential Model</i>	$p = 0.1$	$\omega_1 = 0.1$
	$p = 0.3$	$\omega_1 = 0.2$
	$p = 0.5$	$\omega_1 = 0.3$
	$p = 0.7$	$\omega_1 = 0.4$
	$p = 0.9$	$\omega_1 = 0.5$



(그림 1) 각 모형에 대한 추정된 평균값

## 5. 결 론

소프트웨어 신뢰성은 개발의 최종 단계에 있는 테스트 공정이나 실제 사용단계에 있어서 소프트웨어 내에 존재하는 고장 수나 고장 발생 시간에 의해서 효과적으로 평가할 수 있는 상황으로 그 평가 기술이 중요하게 된다. 따라서 소프트웨어 개발의 테스트 공정이나 실제 사용단계에 있어서 고장 발생 환경이나 고장 발생 현상을 수리적으로 모형화가 가능하면 평가를 할 수 있다. 테스트 시간이나 혹은 실행 시간, 발생된 고장 수와 고장 발생시간과의 관계를 효율적으로 관리함으로서 소프트웨어 신뢰도를 성장 시킬 수 있다. 이러한 과정을 소프트웨어 성장 과정이라고 볼 수 있다.



(그림 2) 각 모형에 대한 신뢰도

본 논문에서는 Minitab(version 14) 통계 패키지에 있는 와이블분포(형상모수가 0.5이고 척도모수가 1)에서 발생시킨 30개의 난수를 이용한 모의 실험 고장 간격시간으로 구성된 자료를 이용하여 기존의 모형과 *Hyper-exponential* 신뢰도 모형에 대하여 최우 추정법을 이용하여 모수 추정을 실시하였다. 효율적인 모형 비교를 위한 편차자승합의 결과는  $p = 0.7$ ,  $\omega_1 = 0.4$ ,  $p = 0.9$ ,  $\omega_1 = 0.5$ 에 따른 *Hyper-exponential* 모형이 기존에 잘 알려진 Yamada-Ohba-Osaki 모형이나 Goel-Okumoto보다 우수함을 보이고 있고 신뢰도에 대한 평가도 역시 *Hyper-exponential* 모형이 신뢰도가 높음을 알 수 있다. 따라서 *Hyper-exponential* 분포를 이용한 신뢰도 모형도 이 분야에서 효율적으로 이용 할 수 있는 모형이 됨을 알 수 있었다. 향후 이러한 *Hyper-exponential* 분포를 이용한 베이지안 적 접근 방법이나 수리적인 추정과 검정 부분에 대한 수리적인 접근이 기대 된다.

## 참 고 문 현

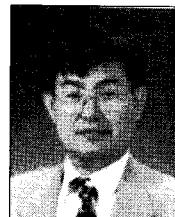
- [1] 김희철, 최유순, 박종구, “어량분포를 이용한 NHPP 소프트웨어 신뢰성장 모형에 관한 연

- 구”, 한국해양정보통신학회논문지, 제10권, 제1호, pp. 7-14, 2006.
- [2] 김희철, “카이제곱 NHPP에 의한 소프트웨어 신뢰성 모형에 관한 연구”, 한국컴퓨터정보학회논문지, 제11권, 제1호, pp. 45-53, 2006.
- [3] S. S. Gokhale and K. S. Trivedi, “A time/structure based software reliability model”, Annals of Software Engineering, Vol. 8, pp. 85-121, 1999.
- [4] J. F. Lawless, “Statistical Models and Methods for Lifetime Data”, John Wiley & Sons, New York, 1981.
- [5] H. Pham, L. Nordmann, and X. Zhang, “A General Imperfect-Software-Debugging Model with S-Shaped Fault-Detection Rate”, IEEE Trans. on reliability, Vol. 48, No 2, pp. 169-175, 1999.
- [6] A. L. Goel and K. Okumoto, “Time-Dependent Error-Detection Rate Models for Software Reliability and Other Performance Measures”, IEEE Trans. on Reliability, Vol. R-28, No. 3, pp. 206-211, Aug. 1979.
- [7] S. Yamada, M. Ohba and S. Osaki, “S-Shaped Reliability Growth Modeling for Software Error Detection”, IEEE Trans. on Reliability, Vol. R-32, No. 5, pp. 475-485, Dec. 1983.
- [8] K. Kanoun and J. C. Laprie, Handbook of Software Reliability Engineering, M. R. Lyu, Editor, chapter Trend Analysis, pp. 401-437.
- [9] S. Brocklehurst and B. Littlewood, “Handbook of Software Reliability Engineering”, M. R. Lyu, Editor, chapter Techniques for Prediction Analysis and Recalibration, pp. 119-166, 1996.
- [10] [www.answers.com/topic/hyper-expontial-distribution](http://www.answers.com/topic/hyper-expontial-distribution).



김희철

1992년 동국대학교 통계학과  
(이학석사)  
1998년 동국대학교 통계학과  
(이학박사)  
2005년 ~ 현재 남서울대학교  
산업경영공학과  
전임강사



신현철

2002년 원광대학교 컴퓨터공학과  
(공학박사)  
2004년 (주)아이비루션 자문위원  
1994년 ~ 현재 백석문화대학  
컴퓨터정보학부 부교수