

일반화 감마 분포를 이용한 NHPP 소프트웨어 신뢰도 모형에 관한 연구

김희철*

NHPP Software Reliability Model based on Generalized Gamma Distribution

Kim-Hee Cheul *

요 약

유한고장 속성을 가진 비동질적인 포아송 과정에 기초한 모형들에서 잔존 결함 1개당 고장 발생률은 일반적으로 상수, 혹은 단조증가 및 단조 감소 추세를 가지고 있다. 본 논문에서는 잔존 결함 1개당 고장 발생률이 단조증가(혹은, 단조감소) 속성을 가진 일반화 감마분포를 이용한 신뢰성 모형을 제안하였다. 일반화 감마분포를 이용한 유한 속성 비동질적 모형에 대한 모수추정은 고장 간격시간으로 구성된 실측자료를 이용하여 모수 추정을 수행하였다. 일반적 감마분포 형상모수의 제안을 위하여 특수한 형태를 적용하였다. 본 논문에서는 기존 모형의 분포를 적용하고 추가적인 소프트웨어 고장 해석을 위하여 감마 및 와이블 분포를 이용하였다. 일반화 감마 분포모형의 고장자료분석을 위하여 산술적 및 라플라스 검정, 적합도 검정, 편의 검정 등을 이용하였다.

Abstract

Finite failure NHPP models presented in the literature exhibit either constant, monotonic increasing or monotonic decreasing failure occurrence rates per fault. This paper proposes reliability model using the generalized gamma distribution, which can capture the monotonic increasing(or, monotonic decreasing) nature of the failure occurrence rate per fault. Equations to estimate the parameters of the generalized gamma finite failure NHPP model based on failure data collected in the form of interfailure times are developed. For the sake of proposing shape parameter of the generalized gamma distribution, used to the special pattern. Data set, where the underlying failure process could not be adequately described by the knowing models, which motivated the development of the gamma or Weibull model. Analysis of failure data set for the generalized gamma model, using arithmetic and Laplace trend tests, goodness-of-fit test, bias tests is presented.

- ▶ Keyword : 소프트웨어 신뢰도 모형(Software Reliability Model), 적합도 검정(Goodness-of-fit Test), 비동질적인 포아송 과정(Nonhomogeneous Poisson Process), 어랑분포(Erlang Distribution), 편의 검정(Bias Tests), 소프트웨어 신뢰도(Software Reliability), 편차자승합(Sum of the Squared Errors).

• 제1저자 : 김희철

• 접수일 : 2005.10.10, 심사완료일 : 2005.11.07

* 남서울대학교 산업경영공학과 전임강사

※ 이 논문은 2005년도 남서울대학교 학술연구조성비 지원에 의하여 연구되었음.

I. 서론

현대사회에 있어서 인간이 의존하고 있는 시스템의 대부분은 소프트웨어 시스템(System)을 포함하고 있으며 이러한 시스템의 기능이 확대되어 소프트웨어가 대규모화, 복잡화, 다양화하고 수요량도 급속히 증대하고 있다. 그러나 시스템이 고장(Failure)이 발생하면 고장의 원인이 되는 결함(Fault)을 찾아 제거하는 디버깅 과정을 거치면서 소프트웨어에 잔존하는 결함의 수는 점차 감소하며 고장이 발생하는 시간 간격이 점차 증가함으로 소프트웨어 신뢰성은 성장한다.

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

소프트웨어 고장시간은 수명자료가 된다. 따라서 비음(Nonnegative)의 값을 가지기 때문에 수명분포가 한정되어 있다. 시간이 지날 수록 비증가 패턴을 가지는 지수분포(Exponential)나, 혹은 와이불(Weibull), 로그정규(Lognormal), 감마분포(Gamma distribution)등을 사용하여 모형을 적용하는 것이 일반적이다. 이러한 분포들은 일반화 감마 분포(Generalized gamma distribution)의 특수한 경우로 알려져 있다[1].

본 연구에서는 이러한 일반화 감마 분포의 특수한 패턴에 대한 NHPP 모형에 대한 신뢰성 척도를 추정하고 이를 바탕으로 신뢰도를 추정하여 그 특성과 효율성에 대하여 고찰 하고자 한다.

II. 유한 고장 NHPP 모형

신뢰도에서 관측시간 $(0, t]$ 사이에 발견된 고장수 $N(t)$ 를 모형화 하는데 비동질적 포아송 과정(Non-homogeneous poisson 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 가정한다. 이 범주에서 지금까지 알려진 모형들은 Goel-Okumoto 모형, Weibull 모형 그리고 Cox-Lewis 모형등이 있는데 이 모형들에 대한 강도함수는 각각 시간에 관한 부분(Fraction) 함수, 멱(Power) 함수, 대수선형(Log-linear) 함수를 가정하였다[1,2].

NHPP 모형에서 평균값 함수 $m(t)$ (Mean value function)와 강도 함수 $\lambda(t)$ 는 다음과 같은 관계로 표현할 수 있다[3].

$$m(t) = \int_0^t \lambda(s) ds \dots\dots\dots (2.1)$$

$$\frac{d m(t)}{dt} = \lambda(t) \dots\dots\dots (2.2)$$

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

$$P\{N(t) = n\} = \frac{[m(t)]^n \cdot e^{-m(t)}}{n!}, n = 0, 1, 2, \dots, \infty \quad (2.3)$$

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[4]. 유한 고장 NHPP 모형들은 충분한 테스트 시간이 주어지면

결합들(Faults)의 기대값이 유한 값 $\left(\lim_{t \rightarrow \infty} m(t) = \theta < \infty \right)$ 을 가지고 반면에 무한 고장 NHPP 모형들은 무한 값을 가진다고 가정 된다. 유한 고장 NHPP 모형에서 충분한 테스트 시간이 주어졌을 때 탐색되어 질 수 있는 결합의 기대값을 θ 라고 표현하고 $F(t)$ 를 분포함수라고 표현하면 유한 고장 NHPP모형의 평균값 함수는 다음과 같이 표현 할 수 있다[3].

$$m(t) = \theta F(t) \dots\dots\dots (2.4)$$

(2.4) 식으로 부터 순간고장 강도함수

(Instantaneous failure intensity) $\lambda(t)$ 는 다음과 같이 유도된다.

$$\lambda(t) = \theta F'(t) \dots\dots\dots (2.5)$$

(2.5) 식을 다음과 같이 변형하여 표기 할 수도 있다.

$$\lambda(t) = [\theta - m(t)] \frac{F'(t)}{1 - F(t)} = [\theta - m(t)] h(t) \dots\dots\dots (2.6)$$

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.

따라서 $[\theta - m(t)]$ 의 값은 시점 t 에 대한 단조 비증가 함수(Monotonically nonincreasing function)가 된다. 즉, 시간이 지남에 따라 결합들이 탐색되어 제거되기 때문에 감소성을 가진다. 따라서 $\lambda(t)$ 는 $h(t)$ 의 값에 따라 달라지며 상수, 증가, 감소 혹은 증가하다가 감소하는 패턴을 가질 수 있다. 이 분야의 기본적 모형인 Goel-Okumoto 모형은 $h(t)$ 가 정수 패턴을 가짐으로서 시점 t 에 독립이고 잘 알려진 S 모형(Yamada, Ohba-Osaki 모형)은 증가패턴을 가진다[3].

III. 일반화 감마분포를 이용한 신뢰성 모형

본 절에서는 일반화 감마분포(1.4)를 이용한 신뢰성 모형을 설명 하고자 한다. 우선 일반화 감마분포의 확률밀도 함수는 다음과 같다.

$$f_{GG}(t|k, \alpha, \beta) = \frac{1}{\Gamma(k)} \beta^k \alpha t^{\alpha k - 1} e^{-\beta t^\alpha}, \dots\dots (3.1)$$

$t > 0, \alpha, \beta, k > 0.$

따라서 분포함수는 다음과 같이 나타낼 수 있다.

$$F_{GG}(t|\alpha, \beta) = I_k(\beta^\alpha) \dots\dots\dots (3.2)$$

단, $I_k(s) = \frac{1}{\Gamma(k)} \int_0^s x^{k-1} e^{-x} dx.$

따라서 위험함수는 다음과 같이 정의된다.

$$h_{GG}(t|k, \alpha, \beta) = \frac{f_{GG}(t|k, \alpha, \beta)}{1 - F_{GG}(t|k, \alpha, \beta)} \dots\dots\dots (3.3)$$

위 일반화 감마분포는 3가지 특성의 분포로 구분해 볼 수 있다. 형상모수(Shape parameter)가 $\alpha=1$ 인 경우는 감마분포 $Gamma(k, \beta)$ 가 되고 $k=1$ 인 경우에는 와이불분포 $Weibull(\alpha, \beta)$ 가 되고 $k \rightarrow \infty$ 일 경우에는 로그 정규분포(극한분포)가 된다[1].

본 연구에서는 감마분포 $Gamma(k, \beta)$ 일 경우와 와이불분포 $Weibull(\alpha, \beta)$ 일 경우에 대해서만 고려 하고자 한다.

3.1 수명 분포가 감마분포 $Gamma(k, \beta)$ 를 적용할 경우

감마분포의 확률밀도함수

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

이 고 분포함수는 k 값이 양의 정수일 때

$$F_{Ga}(t | k, \beta) = 1 - e^{-\beta t} \left(\sum_{i=0}^{k-1} \frac{(\beta t)^i}{i!} \right) \text{가 된다(4.5).}$$

따라서 (2.4)식과 (2.5)에 관련하여 소프트웨어 결함(Fault)당 고장(Failure) 발생률이 일정하거나 증가 혹은 감소하는 특징을 가지는 모형이 감마 모형이 되고 유한 고장 NHPP모형의 평균값 함수 $m(t)$ 와 강도함수 $\lambda(t)$ 는 다음과 같이 표현 할 수 있다.

$$m(t) = \theta F_{Ga}(t | k, \beta) = \theta \left[1 - e^{-\beta t} \left(\sum_{i=0}^{k-1} \frac{(\beta t)^i}{i!} \right) \right] \dots\dots\dots (3.4)$$

$$\lambda(t) = \theta f_{Ga}(t | k, \beta) = \theta \frac{\beta^k}{\Gamma(k)} t^{k-1} e^{-\beta t} \quad (3.5)$$

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 (3.4)식과 (3.5)식에서 형상모수 $k=1$ 인 경우와 같이 된다.

결국 이 모형에 대한 유한 고장 NHPP모형의 평균값 함수 $m(t)$ 와 강도함수 $\lambda(t)$ 는 다음과 같이 표현 할 수 있다.

$$m(t) = \theta F_{Exp}(t | \beta) = \theta (1 - e^{-\beta t}) \dots\dots\dots (3.6)$$

$$\lambda(t) = \theta f_{Exp}(t | \beta) = \beta e^{-\beta t} \dots\dots\dots (3.7)$$

Goel-Okumoto 모형은 소프트웨어 결함(Fault)당 고장(Failure) 발생률이 일정한 형태를 가짐을 가정 하였는데

실제적으로 테스트 초기에 발견되는 결함과 그 이후에 발견되는 결함은 본질적으로 다르기 때문에 실제 소프트웨어 고장수 모형에 적용하기가 적합하지 않은 경우가 있다. 이 경우에 적합하도록 개선한 모형이 잘 알려진 S 모형(Yamada, Ohba-Osaki 모형(7))이다. 따라서 이 모형은 수명 분포를 감마분포중에서 형상모수 $k=2$ 인 분포를 말한다. 따라서 이 모형에 대한 유한 고장 NHPP모형의 평균값 함수 $m(t)$ 와 강도함수 $\lambda(t)$ 는 다음과 같이 표현 할 수 있다.

$$m(t) = \theta F_{Ga(2,\beta)}(t | 2, \beta) = \theta [1 - e^{-\beta t} (1 + \beta t)] \dots\dots\dots (3.8)$$

$$\lambda(t) = \theta f_{Ga(2,\beta)}(t | 2, \beta) = \theta \beta^2 t e^{-\beta t} \dots\dots\dots (3.9)$$

따라서 감마분포의 특수한 경우의 분포를 이용하여 여러 가지 모형을 고려 해 볼 수 있다.

여기에서는 어랑분포(Erlang distribution)(4)를 고려하고자 한다. 이 어랑 분포는 감마분포의 특수한 경우로서 형상모수 k 가 정수인 경우를 의미한다. 본 연구에서는 $k=3$ 인 경우를 적용 하고자 한다.

3.2 수명 분포가 와이불분포 $Weibull(\alpha, \beta)$ 일 경우 와이불 확률밀도함수

$$f_{Wei}(t | \alpha, \beta) = \beta \alpha t^{\alpha-1} e^{-\beta t^\alpha} \quad (\alpha, \beta > 0, t > 0)$$

이 고 분포함수는 $F_{Wei}(t | \alpha, \beta) = 1 - e^{-\beta t^\alpha}$ 가 된다(4). 따라서 유한 고장 NHPP모형의 평균값 함수 $m(t)$ 와 강도함수 $\lambda(t)$ 는 다음과 같이 표현 할 수 있다.

$$m(t) = \theta (1 - e^{-\beta t^\alpha}), \lambda(t) = \theta \beta \alpha t^{\alpha-1} e^{-\beta t^\alpha} \dots\dots\dots (3.10)$$

따라서 와이불 분포의 특수한 경우의 분포를 이용하여 여러 가지 모형을 고려 해 볼 수 있다. 여기에서는 랄리 분포(Rayleigh distribution)(4)를 고려하고자 한다. 이 랄리 분포는 와이불분포의 특수한 경우로서 인 경우 $Weibull(\alpha=2, \beta)$ 과 같기 때

문에 NHPP모형의 평균값 함수 $m(t)$ 와 강도함수 $\lambda(t)$ 는 다음과 같이 표현 할 수 있다.

$$m(t) = \theta (1 - e^{-\beta t^2}), \lambda(t) = 1/2 \theta \beta t e^{-\beta t^2} \dots\dots\dots (3.11)$$

본 연구에서는 $\alpha=3$ 인 경우를 적용 하고자 한다.

IV. 신뢰성 모형에 대한 모수 추정

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

$$x_k = \sum_{i=1}^k t_k \quad (k=1,2,\dots,n; 0 \leq x_1 \leq x_2 \leq \dots \leq x_n) \dots\dots\dots (4.1)$$

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 구성된다. 이 시간 절단 모형에서의 우도함수는 다음과 같이 알려져 있다[1,2].

$$L_{NHPP}(\theta, \beta | D_t) = \left(\prod_{i=1}^n \lambda(x_i) \right) \exp(-m(t)) \dots (4.2)$$

단, θ, β 은 미지의 모수들을 의미하고 우도함수 (4.2) 식에서 t 을 x_n 으로 대체하면 유사한 형태의 고장 절단 모형의 우도함수가 된다[2]. (4.2식과 (3.1), (3.2)식을 연관하면 일반화 감마 모형에 대한 우도함수는 다음과 같은 형태로 표현 할 수 있다.

$$L_{NHPP}(\theta, k, \alpha, \beta | D_t) = \left(\prod_{i=1}^n \theta f_{GG}(x_i) \right) \exp(-\theta F_{GG}(x_n)) \dots\dots\dots (4.3)$$

따라서 최우추정법(MLE)을 이용하기 위한 *Gamma*(k, β) 모형 로그우도함수(Log likelihood function)는 다음과 같이 유도된다.

$$\begin{aligned} \ln L(\theta, k, \beta | D_{x_n}) & \dots\dots\dots (4.4) \\ &= n \ln \theta - n \ln \Gamma(k) + nk \ln \beta + (k-1) \sum_{i=1}^n \ln x_i \\ & \quad - \beta \sum_{i=1}^n x_i - \theta + \theta e^{-\beta x_n} \left(\sum_{j=0}^{k-1} \frac{(\beta x_n)^j}{j!} \right) \end{aligned}$$

형상 모수 k 값은 상수(알고 있다고)라고 가정 했을 때 최우추정법을 이용하기 위하여 (4.4)식을 θ 와 β 에 대하여 편미분을 하면 다음과 같은 식을 유도 할 수 있다.

$$\frac{n}{\theta} = 1 - \exp(-\beta x_n) \sum_{i=0}^{k-1} \frac{(\beta x_n)^i}{i!} \dots\dots\dots (4.5)$$

$$\frac{kn}{\beta} = \sum_{i=1}^n x_i - \frac{\theta \left[\theta \exp(-\beta x_n) \sum_{i=0}^{k-1} \frac{(\beta x_n)^i}{i!} \right]}{\theta \beta} \dots (4.6)$$

(4.5) 식과 (4.6) 식을 비선형 연립방정식(수치해석적 방법)을 이용하여 풀면 최우추정치 $\hat{\beta}_{MLE}$ 와 $\hat{\theta}_{MLE}$ 의 값을 구 할 수 있다. (4.5) 식과 (4.6) 식에서 $k=1$ 이면 Goel-Okumoto 모형이되고 $k=2$ 은 Yamada, Ohba-Osaki 모형이 된다. 그리고 k 가 정수이면 어랑(Erlang) 모형이라고 한다. 본 연구에서는 앞의 모형과 비교하여 확장 개념으로 Erlang ($k=3$)인 모형을 고려하고자 한다.

그리고 최우추정법(MLE)을 이용하기 위한 *Weibull*(α, β) 모형 로그우도함수(Log likelihood function)는 다음과 같이 유도된다.

$$\begin{aligned} \ln L(\theta, \alpha, \beta | D_{x_n}) & \dots\dots\dots (4.7) \\ &= n \ln \theta + n \ln \beta + n \ln \alpha \\ & \quad + (\alpha-1) \sum_{i=1}^n \ln x_i - \beta \sum_{i=1}^n x_i^\alpha - \theta + \theta e^{-\beta x_n^\alpha} \end{aligned}$$

형상 모수 α 값은 상수(알고 있다고)라고 가정 했을 때 최우추정법을 이용하기 위하여 (4.7)식을 θ 와 β 에 대하여 편미분을 하면 다음과 같은 식을 유도 할 수 있다.

$$\frac{n}{\theta} = 1 - \exp(-\beta x_n^\alpha), \frac{n}{\beta} = \sum_{i=1}^n x_i^\alpha + x_n^\alpha \theta e^{-\beta x_n^\alpha} \quad (4.8)$$

(4.8)식을 비선형 연립방정식(수치해석적 방법)을 이용하여 풀면 최우추정치 $\hat{\beta}_{ME}$ 와 $\hat{\theta}_{ME}$ 의 값을 구 할 수 있다. 여기에서는 랄리 분포(Rayleigh distribution)[4]를 고려하고자 한다. 이 랄리 분포는 와이불분포의 특수한 경우로서 Weibull($\alpha=2, \beta$)인 경우와 같기 때문에 (4.8)식을 이용하면 된다.

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 될이 알려져 있다[6,8].

$$\hat{R}(t | x_n) = e^{-\int_{x_n}^{x_n+t} \lambda(\tau) d\tau} = \exp\{-\{m(t+x_n) - m(x_n)\}\} \quad (4.9)$$

V. 소프트웨어 고장 자료 분석

이 장에서 Michael R.Lyu(〈표 1〉, [9])가 인용한 고장 간격시간 자료(Failure interval time data)를 가지고 일 반화 감마모형에 대한 신뢰도 모형을 분석하고자 한다.

이 고장 간격 데이터(70.7 시간(hours) 단위에서 고장이 24번 일어남)을 이용하였고 〈표 1〉에 자료가 나열 되어 있다. 적용 하고자 하는 신뢰모형들을 분석하기 위하여 우선 자료에 대한 추세 검정이 선행 되어야 한다[3,10]

추세 분석에는 산술평균 검정(Arithmetic mean test)과 라플라스 추세 검정(Laplace trend test)등이 있다. 이

검정을 실시한 결과 그림인 (그림 1)에서 산술평균 검정결과 고장수가 증가함에 따라 산술 평균이 대체적으로 증가 추세를 보이고 있으므로 신뢰성장(Reliability growth)이 되고 있음을 나타내고 있고 라플라스 추세 검정의 결과도 라플라스 요인(Factor)이 음수로서 감소하기 때문에 역시 신뢰성장(Reliability growth)이 되고 있음을 나타내고 있다. 따라서 이 자료를 가지고 신뢰성장 모형을 제시하는 것이 효율적 임을 시사하고 있다.

소프트웨어 신뢰성 모형의 모수 추정은 최우추정법을 이용하였고 비선형 방정식의 계산방법은 수치해석적 기본 방법인 이분법(Bisection method)을 사용하였다. 이러한 계산은 초기값을 10^{-6} 와 10 을, 허용 한계(Tolerance for width of interval)는 10^{-10} 을 주고 수렴성을 확인 하면서 충분한 반복 횟수인 100번을 C-언어를 이용하여 모수 추정을 수행한 모수의 추정값들의 결과는 〈표 2〉에 요약되었다.

모형 선택의 하나의 방법으로 편차자승합(SSE [7])을 이용할 수 있는데 이 편차자승합이 작으면 상대적으로 효율적인 모형이 된다. 주어진 자료를 이용하여 제시된 모형들에 대한 편차자승합의 값은 〈표 3〉에 요약되었다. 이 표에서 척도모수가 3인($k=3, \alpha=3$)인 어랑 모형이나 와이불 모형이 이 분야에서 기존에 알려진 모형인 Yamada-Ohba-Osaki 모형이나 Goel-Okumotoand 모형에 비해 상대적으로 효율적 모형으로 나타나고 있다.

표 1. 고장 간격 자료
Table 1. Failure interval time

Failure number	Failure Time(hour)	Failure Interval(hour)
1	0.5	0.5
2	1.7	1.2
3	4.5	2.8
4	7.2	2.7
5	10	2.8
6	13	3
7	14.8	1.8
8	15.7	0.9
9	17.1	1.4
10	20.6	3.5
11	24	3.4
12	25.2	1.2
13	26.1	0.9
14	27.8	1.7
15	29.2	1.4
16	31.9	2.7
17	35.1	3.2
18	37.6	2.5
19	39.6	2
20	44.1	4.5
21	47.6	3.5
22	52.8	5.2
23	60	7.2
24	70.7	10.7

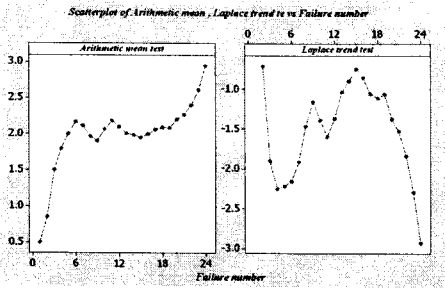


그림 1. 산술평균 검정과 리플라스 추세 검정
Figure 1. Arithmetic mean test and Laplace trend test

예측 오류(prediction error)의 비정상성(nonstationarity)에 대한 측도는 Kolmogorov 거리(distance) [3,10,11]로 측정되는데 이 거리가 클수록 상대적으로 비정상성을 내포하고 있다. (그림 2)은 S-Plus 소프트웨어 [11]를 이용하여 Kolmogorov 검정에 대한 그림을 보여주고 있고 이 그림에서도 전체적으로 어랑모형이 상대적으로 비정상성 속성이 덜 내포하고 있음을 알 수 있다. <표 4>는 Kolmogorov 거리를 요약한 표로서 모형에 대한 치우침(bias)을 나타내고 있다. 이 표에서도 어랑모형의 형태가 상대적으로 비정상성 속성이 덜 내포하고 있음을 알 수 있다.

표 2. 각 모형의 모수 추정값
Table 2. Estimator of each model

모형	β_{ME}	$\hat{\theta}_{ME}$
Goel-Okumoto Model	$10^{-2} \times 4.6431$	28.3798
Yamada-Ohba-Osaki Model	$10^{-2} \times 9.8904$	26.0845
Erlang Model ($k=3$)	$10^{-2} \times 8.7958$	25.8523
Rayleigh Model	$10^{-3} \times 3.8912$	28.0071
Weibull Model ($\alpha=3$)	$10^{-5} \times 2.5283$	24.0032

표 3. 모형들에 대한 SSE 의 값
Table 3. SSE of each model

Model	Goel-Okumoto Model	Yamada-Ohba-Osaki Model	Erlang Model ($k=3$)	Rayleigh Model	Weibull Model ($\alpha=3$)
SSE	746.6847	433.0076	127.1765	605.7211	288.8676

Kolmogorov-test-plot

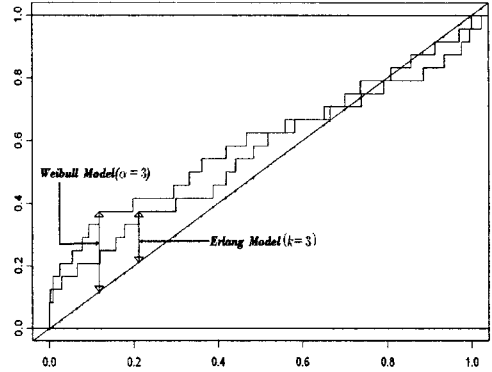


그림 2. 모형에 대한 Kolmogorov 검정 그림
Figure 2. Kolmogorov test of each model

표 4. 모형에 대한 Kolmogorov 거리
Table 4. Kolmogorov distance of each model

Model	Kolmogorov distance
Erlang Model ($k=3$)	0.16316
Weibull Model ($\alpha=3$)	0.256422

어랑 모형과 와이블 모형에 대한 신뢰도를 추정한 결과는 (그림 3)에 요약되었고 이 그림에서도 신뢰도는 처음 시간에는 비슷하다가 시간이 흐를수록 어랑모형이 높아지는 패턴을 볼 수 있다.

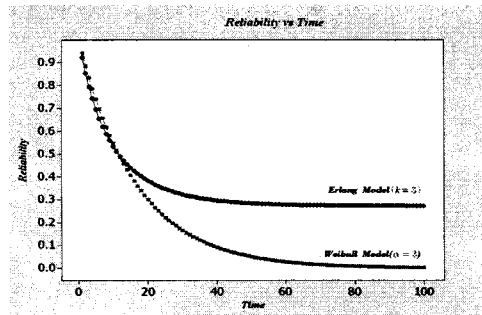


그림 3. 어랑모형과 와이블 모형의 신뢰도
Figure 3. Reliability of Erlang model and Weibull model

어랑 모형에서 앞으로의 신뢰도 예측관계를 위하여 실제 고장수가 20, 22, 24 이후의 신뢰도를 추정한 결과는 (그림 4)에 나열 되었다. 이 그림에서는 임무시간(Mission

time)에 따른 신뢰도가 비증가 추세를 보이고 있으며 실제 고장수가 20 보다는 24이후의 신뢰도가 높음을 알 수 있다. 즉, 고장수가 많아지면 남아있는 결함의 수가 적기 때문이다.

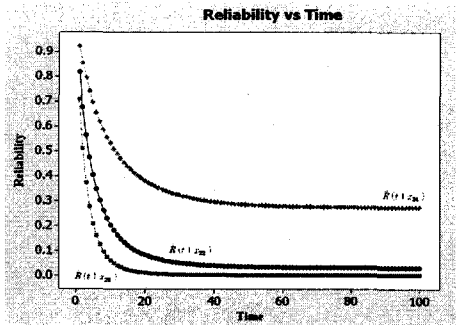


그림 4. 어람모형에서의 임무시간에 따른 신뢰도
Figure 4. Reliability of mission time Arlang model

VI. 결론

소프트웨어 신뢰성은 개발의 최종단계에 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 실제사용단계에 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 관리함으로써 소프트웨어 신뢰도를 성장 시킬 수 있다. 이러한 과정을 소프트웨어 성장 과정이라고 볼 수 있다.

본 논문에서는 잔존 결함 1개당 고장 발생률이 증가 혹은 감소 추세를 가진 일반화 감마분포의 특수한 경우를 이용한 일반화 감마 모형을 제안하였다. 고장 간격시간으로 구성된 자료를 이용하여 형상모수가 알고 있을 때 표준 감마 형태와 와이불 형태 및 기존의 모형에 대하여 최우 추정법을 이용하여 모수 추정을 실시하였다.

효율적 모형을 판단하는 하나의 기준인 편차지수함을 고장수가 24 인 자료를 이용하여 계산 한 결과는 기존의 모형보다 어람 모형이나 와이불 모형이 효율적 모형이 됨을

알 수 있었고 또, 비정상의 척도인 콜모고로프 거리의 결과도 어람 모형이나 와이불 모형이 상대적으로 정상성에 가까운 모형으로 간주 할 수 있다.

어람 모형과 와이불 모형에 대한 신뢰도를 추정한 결과는 신뢰도는 처음시간에는 비슷하다가 시간이 흐를 수록 어람모형이 높아지는 패턴을 볼 수 있었고 어람 모형에서 임무시간에 따른 신뢰도가 비증가 추세를 보이고 있으며 고장수가 많아지면 남아있는 결함의 수가 적기 때문에 실제고장수가 20 보다는 24이후의 신뢰도가 높음을 알 수 있다.

분석된 자료에 대한 평가에서도 산술평균 검정과 리플라스 추세 검정을 실시한 결과도 신뢰성장이 되고 있음을 나타내고 있다.

참고문헌

- [1] J. F. Lawless. Statistical Models and Methods for Lifetime Data. John Wiley & Sons, New York, 1981.
- [2] L. Kuo and T. Y. Yang. "Bayesian Computation of Software Reliability". Journal of the American Statistical Association, Vol.91, pp.763-773, 1996.
- [3] S. S. Gokhale and K. S. Trivedi. "A time/structure based software reliability model". Annals of Software Engineering, 8, 85-121. 1999
- [4] V. K. Rohatgi. Statistical inference, pages 398-416. JOHN WILEY & SONS, INC., New York, 1984.
- [5] [http://www.answers.com/Weibull%20Distribution%20\(site visited 2005-7-15\)](http://www.answers.com/Weibull%20Distribution%20(site%20visited%202005-7-15))
- [6] A. L. Goel and K. Okumoto. "Time-Dependent Error-Detection Rate Models for Software Reliability and Other Performance Measures". IEEE Trans. on Reliability, R-28(3):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, R-32(5):475-485, Dec. 1983.

- [8] H. Pham and 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, 1999.
- [9] M. R. Lyu. Handbook of Software Reliability Engineering, M.R.Lyu, Editor, chapter Introduction, pp.3-25. MacGraw-Hill, New York, 1996
- [10] K. Kanoun and J. C. Laprie. Handbook of Software Reliability Engineering, M.R.Lyu, Editor, chapter Trend Analysis, pp.401-437. MacGraw-Hill, New York, 1996.
- [11] S. Selvin. Modern Applied Biostatistical Methods Using S-Plus, pages 141-184. Oxford University Press, New York, 1998.
- [12] 김홍진, "소프트웨어 재사용을 위한 소프트웨어 칩 표 현식에 관한 연구". 한국컴퓨터정보학회논문지, v.006, n.004, pp.12-20, 2001, 12
- [13] 정창신, 정순기 "소프트웨어 자동 테스트 도구의 발전 로드맵분석". 한국컴퓨터정보학회논문지, v.9, n.1, pp.17-23, 2004. 03

저자 소개



김희철

1992년 동국대학교 통계학과 석사
 1998년 동국대학교 통계학과 박사
 2000년 3월~2004년 2월 송호대학
 정보산업계열 조교수
 2005년 3월~현재 남서울대학교
 산업경영공학과 전임강사
 <관심분야> 소프트웨어 신뢰성공학,
 웹프로그래밍, 전산통계,
 인터넷비즈니스