韓國 컴퓨터情報學會 論文誌
第11卷 第1號, 2006. 3.

2006-11-1-1-6

# 카이제곱 NHPP에 의한 소프트웨어 신뢰성 모형에 관한 연구

김 희 철*

# The Study for NHPP Software Reliability Model based on Chi-Square Distribution

Hee-Cheul Kim*

## 요 약

유한고장수를 가진 비동질적인 포아송 과정에 기초한 모형들에서 잔존 결함 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 잔존 결함 1개당 고장 발생률이 증가추세를 가진 카이제곱 분포를 이용한 카이제곱 모형을 제안하였다. 고장 간격시간으로 구성된 자료를 이용한 모수추정 방법은 최우추정법과 일반적인 수치해석 방법인 이분법을 사용하여 모수 추정을 실시하고 효율적인 모형 선택은 편차자승합, AIC 통계량 및 콜모고로프 거리를 적용하여 모형들에 대한 효율성 입증방법을 설명하였다. 소프트웨어 고장 자료 분석에서는 카이제곱 모형에 대한 자유도를 형상모수의 척도로 간주하여 고장수가 비교적 큰 실측 자료(고장수가 86)인 Allen P.Nikora 와 Michael R.Lyu가 인용한 SYS2 자료을 통하여 분석하였다. 이 자료들에서 카이제곱 모형의 비교를 위하여 산술적 및 라플라스 검정, Kolmogorov 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. In this paper, Goel-Okumoto and Yamada-Ohba-Osaki model was reviewed, proposes the $\chi^2$ reliability model, which can capture the increasing nature of the failure occurrence rate per fault. Algorithm to estimate the parameters used to maximum likelihood estimator and bisection method, model selection based on SSE, AIC statistics and Kolmogorov distance, for the sake of efficient model, was employed. Analysis of failure using real data set, SYS2(Allen P.Nikora and Michael R.Lyu), for the sake of proposing shape parameter of the $\chi^2$ distribution using the degree of freedom, was employed. This analysis of failure data compared with the $\chi^2$ model and the existing model using arithmetic and Laplace trend tests, Kolmogorov test is presented.

▸ Keyword : 소프트웨어 신뢰도 모형(Software Reliability Model), 비동질적인 포아송 과정(Nonhomogeneous Poisson Process), 카이제곱분포($\chi^2$ Distribution), 편의 검정(Bias Tests), 소프트웨어 신뢰도(Software Reliability), 편차자승합(Sum of the Squared Errors), 콜모고로프 거리(Kolmogorov distance)

• 제1저자 : 김희철
• 접수일 : 2005.12.06, 심사완료일 : 2006.02.24
* 남서울대학교 산업경영공학과 전임강사

# I. 서 론

소프트웨어 테스팅 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 이 분야에서는 주로 지수분포, 와이블분포, 감마분포 등 일반화 감마 분포가 많이 사용되어 왔다. 그러나 이러한 감마분포 외에도 통계자료 해석시 자료를 대수변환 이 후에 정규분포로 처리하면 되는 대수정규분포(Lognormal distribution), 이 분포와 유사한 대수로지스틱 분포(Loglogistic distribution)도 이 분야에 사용이 가능하다[1]. 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 이러한 카이제곱 분포를 적용하여 NHPP 모형에 대한 신뢰성 척도를 추정하고 이를 바탕으로 자유도에 따른 모형의 효율성과 그 특성을 알아보고자한다.

본 논문의 2장에서는 관련연구로서 유한 고장 NHPP 모형에 대하여 서술하였고 3장에서는 카이제곱 신뢰성 모형에 대하여 알아보았고 4장에서는 최우추정법을 이용한 모수추정에 대하여 설명하고 5장에서는 실측 소프트웨어 고장자료인 Allen P.Nikora 와 Michael R.Lyu 가 인용한 SYS2 자료를 이용하여 각 모형에 대한 모수추정 및 모형비교를 실시하였으며 마지막으로 5장에서는 결론을 서술 하였다.

# II. 관련 연구

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[3]. 이 범주에서 지금까지 알려진 모형들은 Goel-Okumoto 모형, Weibull 모형 그리고 Cox-Lewis 모형등이 있는데 이 모형들에 대한 강도함수는 각 각 시간에 의존한 함수, 멱(Power) 함수, 대수선형(Log-linear) 함수를 가정하였다[2].

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

$$m(t) = \int_0^t \lambda(s)ds \ , \quad \frac{dm(t)}{dt)} = \lambda(t) \quad \cdots\cdots\cdots\cdots (2.1)$$

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

$$P\{N(t) = n\} = \frac{[m(t)]^n \cdot e^{-m(t)}}{n!}, \ n = 0,1,2,\ldots\infty \cdot (2.2)$$

이처럼 시간영역 모형(Time domain models)들은 NHPP 에 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[3].

유한 고장(Finite failure) NHPP 모형들은 충분한 테스트 시간이 주어지면 결함들(Faults)의 기대값이 유한 값 $\left( \lim_{t\to\infty} m(t) = \theta < \infty \right)$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 결함의 기대값을 $\theta$라고 표현하고 $F(t)$를 분포함수라고 표현하면 유한 고장 NHPP모형의 평균값 함수는 다음과 같이 표현 할 수 있다[2].

$$m(t)=\theta F(t) \quad \cdots\cdots (2.3)$$

(2.3)식으로 부터 순간고장 강도함수(Instantaneous failure intensity function) $\lambda(t)$는 다음과 같이 유도된다.

$$\lambda(t)=\theta F'(t) \quad \cdots\cdots (2.4)$$

(2.4)식을 다음과 같이 변형하여 표기 할 수도 있다.

$$\lambda(t)=[\theta-m(t)]\frac{F'(t)}{1-F(t)}=[\theta-m(t)]\,h(t) \quad \cdots\cdots (2.5)$$

단, $h(t)=\frac{F'(t)}{1-F(t)}$ 는 위험함수(Hazard function, 고장률 함수)으로서 소프트웨어 결함당 고장 발생률을 의미하고 $[\theta-m(t)]$은 $t$ 시점에서 소프트웨어에 남아있는 결함들의 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 수 있다. 이 분야의 기본적 모형인 Goel-Okumoto 모형은 $h(t)$가 정수 패턴을 가짐으로서 시점 $t$에 독립이고 잘 알려진 Yamada, Ohba-Osaki 모형은 단조 비감소 패턴을 가진다[1].

## III. 카이제곱분포를 이용한 신뢰성 모형

이 장에서 2변량 카이제곱분포($\sigma^2\chi^2$ distribution)[4]를 이용한 신뢰성 모형을 설명하고자 한다. 우선 2변량 카이제곱분포의 확률밀도 함수는 다음과 같다.

$$f_{\chi^2}(t\mid\nu,\sigma^2)=\frac{1}{2^{\nu/2}\,\sigma^{\nu}\,\Gamma(\nu/2)}\,t^{\nu/2-1}\,e^{-t/2\sigma^2} \quad \cdots\cdots (2.6)$$

단, $t>0$, $\nu>0$는 자유도이며 (2.6)식의 분포함수[4]는 다음과 같이 알려져 있다.

$$F_{\chi^2}(t\mid r,\sigma^2)=1-\exp\left(-\frac{t}{2\sigma^2}\right)\sum_{i=0}^{r-1}\frac{(t/2\sigma^2)^i}{i!} \quad \cdots\cdots (2.7)$$

따라서 위험함수는 다음과 같이 정의된다.

$$h_{\chi^2}(t\mid\nu,\sigma^2)=\frac{f_{\chi^2}(t\mid\nu,\sigma^2)}{1-F_{\chi^2}(t\mid r,\sigma^2)} \quad \cdots\cdots (2.8)$$

여기서 양의정수 $r=\nu/2$이고 $\sigma^2$는 척도모수(Scale parameter)를 의미하며 $\nu$는 자유도(Degree of freedom, df) 이면서 형상모수(Shape parameter)이다.

소프트웨어 결함당 고장 발생률이 일정하거나 증가하는 특징을 가지는 모형이 카이제곱 분포가 되고 유한 고장 NHPP모형의 평균값 함수 $m(t)$와 강도함수 $\lambda(t)$는 다음과 같이 표현 할 수 있다.

$$m(t)=\theta F_{\chi^2}(t\mid r,\sigma^2)=\theta\left[1-\exp\left(-\frac{t}{2\sigma^2}\right)\sum_{i=0}^{r-1}\frac{(t/2\sigma^2)^i}{i!}\right] \quad \cdots\cdots (2.9)$$

$$\lambda(t)=\theta f_{\chi^2}(t \mid \nu, \sigma^2)=\theta \frac{1}{2^{\nu/2}\sigma^{\nu}\Gamma(\nu/2)} t^{\nu/2-1} e^{-t/2\sigma^2} \quad \cdots\cdots (2.10)$$

(그림 1) 은 자유도에 따른 카이제곱 모형의 위험함수를 나타낸 그림이다. 이 모형의 모수들은 <표 1>에 있는 고장 간격 데이터을 이용하여 추정된 결과를 이용하였다. 이 그림에서 Goel-Okumoto모형과 자유도가 2인 카이제곱 모형은 시간과 독립적으로 나타나고 있고 Yamada-Ohba-Osaki 모형, 자유도가 4, 6인 카이제곱 모형은 단조 비감소 (Monotonical nondecreasing)형태를 보이고 있음을 확인 할 수 있다.

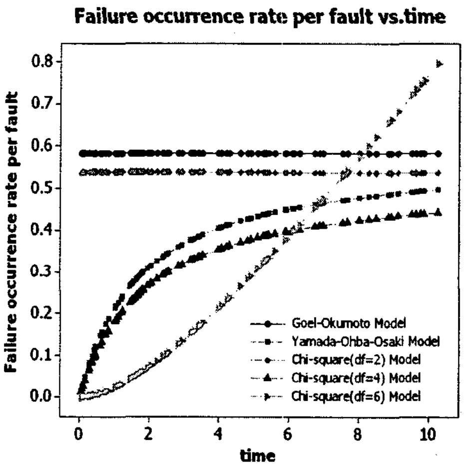


그림 1. 각 모형에 대한 위험함수
Figure 1. Hazard function of each model

## IV. 신뢰성 모형에 대한 모수 추정

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

$$x_k=\sum^{k} t_k \ (k=1,2,\cdots,n \,;\, 0 \leq x_1 \leq x_2 \leq \cdots \leq x_n) \quad \cdots (2.11)$$

이라고 하면 데이터 집합 $D_t$는 $\{n,x_1,x_2\cdots x_n\,;t\}$와 같이 구성된다. $n$번째까지 고장시점이 관찰된 고장 절단 모형일 경우에 데이터 집합 $D_{x_n}$은 $\{x_1,x_2\cdots x_n\}$으로 구성된다.

이 시간 절단 모형에서의 $\Theta$를 모수공간이라고 표시하면 우도함수는 다음과 같이 알려져 있다[2,3].

$$L_{NHPP}(\Theta \mid D_{x_n})=\left(\prod_{i=1}^{n}\lambda(x_i)\right)\exp(-m(x_n)) \quad \cdots\cdots (2.12)$$

### 4.1 Goel-Okumoto모형과 Yamada-Ohba-Osaki 모형

Goel-Okumoto 모형[5]에 대한 평균값 함수는 $m_1(t)=\theta(1-e^{-\beta_1 t})$ $(\theta>0,\beta_1>0)$이라고 알려져 있다.

$t$을 최종 고장시점 $x_n$으로 대치하고 고장 발생률 $\lambda_1(t)=\theta\beta_1 e^{-\beta_1 t}$ 을 이용하면 우도함수는 다음과 같이 표현 할 수 있다.

$$L_{GO}(\theta,\beta_1 \mid D_{x_n})=\left(\prod_{k=1}^{n}\theta\beta_1 e^{-\beta_1 x_k}\right)\exp\left[-\theta(1-e^{-\beta_1 x_n})\right] \cdot \quad \cdots\cdots (2.13)$$

최우추정법(MLE)을 이용하기 위한 로그우도함수를 구하면,

$$\ln L_1=n\ln\theta+n\ln\beta_1-\beta_1\sum_{k=1}^{n}x_k-\theta(1-e^{-\beta_1 x_n}) \quad \cdots\cdots (2.14)$$

으로 표현된다.

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

$$\frac{\partial \ln L_1}{\partial\theta}=\frac{n}{\theta}-1+e^{-\beta_1 x_n}=0,$$

$$\frac{\partial \ln L_1}{\partial\beta_1}=\frac{n}{\beta_1}-\sum_{k=1}^{n}s_k-\theta\, x_n e^{-\beta_1 x_n}=0$$

각 모수에 대한 최우추정량 $\hat{\theta}_{MLE}$ 와 $\hat{\beta}_{1MLE}$ 은 다음식을 만족한다.

$$\frac{n}{\hat{\theta}}=1-\exp\left(-\hat{\beta}_1 x_n\right), \quad (2.15)$$

$$\frac{n}{\hat{\beta}_1}=\sum_{k=1}^{n} x_k+\hat{\theta} x_n \exp\left(-\hat{\beta}_1 x_n\right) \quad (2.16)$$

(2.15)식과 (2.16)식을 수치적으로 풀어 두 모수를 구할 수 있다.

Yamada-Ohba-Osaki 모형[6]에 대한 평균값 함수는 $m_2(t)=\theta\left[1-(1+\beta_2 t)e^{-\beta_2 t}\right]$ $(\theta>0, \beta_2>0)$이라고 알려져 있고 $\lambda_2(t)=\theta\beta_2^2 te^{-\beta_2 t}$ 가 되고 우도함수는 다음과 같다.

$$L_{YOO}(\theta,\beta_2 \mid D_{x_n})=\left(\prod_{k=1}^{n}\theta\beta_2^2 x_k e^{-\beta_2 x_k}\right) \quad (2.17)$$
$$\cdot \exp[-\theta[1-(1+\beta_2 x_n)e^{-\beta_2 x_n}]$$

따라서 로그우도함수를 구하면 다음과 같이 표현된다.

$$\ln L_2 = n\ln\theta+2n\ln\beta_2+\sum_{k=1}^{n}\ln x_k-\beta_2\sum_{k=1}^{n}x_k \quad (2.18)$$
$$-\theta[1-(1+\beta_2 x_n)e^{-\beta_2 x_n}]$$

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

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

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

$$\frac{n}{\hat{\theta}}=1-\exp(-\hat{\beta}_2 x_n)-\hat{\beta}_2 x_n\exp(-\hat{\beta}_2 x_n) \quad (2.19)$$

$$\frac{2n}{\hat{\beta}_2}=\sum_{k=1}^{n}x_k+\hat{\theta}\,\hat{\beta}_2 x_n^2\exp(-\hat{\beta}_2 x_n) \quad (2.20)$$

(2.19)식과 (2.20)식을 수치적으로 풀어 두 모수를 구할 수 있다.

## 4.2 카이제곱 모형

(2.12)식과 (2.9), (2.10)식을 연관하면 카이제곱 모형에 대한 우도함수는 다음과 같은 형태로 표현 할 수 있다.

$$L_{NHPP}(\theta,\nu,\sigma^2 \mid D_t)=\left(\prod_{i=1}^{n}\theta f_{\chi^2}(x_i)\right)\exp\left(-\theta F_{\chi^2}(x_n)\right) \quad (2.21)$$

따라서 최우추정법을 이용하기 위한 카이제곱 모형 로그우도함수는 다음과 같이 유도된다
(단, $r=\nu/2$ )[4].

$$\ln L(\theta,\sigma^2 \mid D_{x_n}) \quad (2.22)$$
$$=n\ln\theta+(\nu/2-1)\sum_{i=1}^{n}\ln x_i-\sum_{i=1}^{n}\frac{x_i}{2\sigma^2}$$
$$-\frac{\nu n}{2}\ln 2-n\ln\Gamma(\nu/2)-\nu n\ln\sigma-\theta$$
$$+\theta\exp\left(-\frac{x_n}{2\sigma^2}\right)\sum_{i=0}^{r-1}\frac{(x_n/2\sigma^2)^i}{i!}$$

형상 모수 $\nu$ 값은 상수(사전에 알고 있는 경우)라고 가정 했을 때 최우추정법을 이용하기 위하여 (2.22)식을 $\theta$ 와 $\sigma^2$ 에 대하여 편미분을 하면 다음과 같은 식을 유도 할 수 있다.

$$\frac{n}{\theta}=1-\exp\left(-\frac{x_n}{2\sigma^2}\right)\sum_{i=0}^{r-1}\frac{(x_n/2\sigma^2)^i}{i!} \quad (2.23)$$

$$\frac{\nu n}{2\sigma^2}= \quad (2.24)$$
$$\frac{\sum_{i=1}^{n}x_i}{2(\sigma^2)^2}-\frac{\partial\left[\theta\exp\left(-\frac{x_n}{2\sigma^2}\right)\sum_{i=0}^{r-1}\frac{(x_n/2\sigma^2)^i}{i!}\right]}{\partial\sigma^2}$$

(2.23) 식과 (2.24) 식을 비선형 연립방정식(수치해석적 방법)을 이용하여 풀면 최우추정치 $\hat{\sigma^2}_{MLE}$와 $\hat{\theta}_{MLE}$ 의 값을 구 할 수 있다.

그러나 카이제곱 분포의 분포함수는 (2.7)식에서 보여 주듯이 자유도 $\nu$가 짝수인 경우에만 닫힌 형태(폐쇄형; Closed form)가 되므로 본 논문에서는 자유도가 짝수인 경우 즉, df=2, 4, 6인 경우를 본 논문에서는 제시하고자 한다.

## V. 소프트웨어 고장 자료 분석

이 장에서 실제적인 자료를 가지고 자유도에 따른 카이제곱 신뢰도 모형을 분석하고자 한다. 고장자료는 Allen P.Nikora 와 Michael R.Lyu 가 인용한 SYS2 자료[7]을 이용하고자 한다.

실제 분석에서는 원래의 자료를 변량변환(Variate transformation) 시킨 고장 간격 데이터($102594 \times 10^{-4}$ 시간(Second) 단위에서 고장이 86번 일어남)을 이용하였고 <표 1>에 자료가 나열 되어 있다.

제시하는 신뢰모형들을 분석하기 위하여 우선 자료에 대한 추세 검정이 선행 되어야 한다[1,8].

추세 분석에는 산술평균 검정(Arithmetic mean test)과 라플라스 추세 검정(Laplace trend test)[7]등이 있다. 이 검정을 실시한 결과 (그림 2)에서 산술평균 검정결과 고장수가 증가함에 따라 산술 평균이 거의 증가 추세를 보이고 있으므로 신뢰성장(Reliability growth)이되고 있음을 나타내고 있고 (그림 3)에 나타난 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[8].

소프트웨어 신뢰성 모형의 모수 추정은 최우추정법을 이용하였고 비선형 방정식의 계산방법은 수치해석적 기본 방법인 이분법(Bisection method)을 사용하였다. 이러한 계산은 초기값을 $10^{-1}$ 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 수행하였다.

자유도에 따른 카이제곱 모형을 포함한 모수의 추정값들의 결과는 <표 2>에 요약되었다.

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

| *Failure number* | *Failure interval(second)* | *Failure time(second)* | *Variate transformation time(second)* |
|---|---|---|---|
| 1 | 479 | 479 | 0.0479 |
| 2 | 266 | 745 | 0.0745 |
| 3 | 277 | 1022 | 0.1022 |
| 4 | 554 | 1576 | 0.1576 |
| 5 | 1034 | 2610 | 0.261 |
| 6 | 249 | 2859 | 0.2859 |
| 7 | 693 | 3552 | 0.3552 |
| 8 | 597 | 4149 | 0.4149 |
| 9 | 117 | 4266 | 0.4266 |
| 10 | 170 | 4436 | 0.4436 |
| 11 | 117 | 4553 | 0.4553 |
| 12 | 1274 | 5827 | 0.5827 |
| 13 | 469 | 6296 | 0.6296 |
| 14 | 1174 | 7470 | 0.747 |
| 15 | 693 | 8163 | 0.8163 |
| 16 | 1908 | 10071 | 1.0071 |
| 17 | 135 | 10206 | 1.0206 |
| 18 | 277 | 10483 | 1.0483 |
| 19 | 596 | 11079 | 1.1079 |
| 20 | 757 | 11836 | 1.1836 |
| 21 | 437 | 12273 | 1.2273 |
| 22 | 2230 | 14503 | 1.4503 |
| 23 | 437 | 14940 | 1.494 |
| 24 | 340 | 15280 | 1.528 |
| 25 | 405 | 15685 | 1.5685 |
| 26 | 535 | 16220 | 1.622 |
| 27 | 277 | 16497 | 1.6497 |
| 28 | 363 | 16860 | 1.686 |
| 29 | 522 | 17382 | 1.7382 |
| 30 | 613 | 17995 | 1.7995 |
| 31 | 277 | 18272 | 1.8272 |
| 32 | 1300 | 19572 | 1.9572 |
| 33 | 821 | 20393 | 2.0393 |
| 34 | 213 | 20606 | 2.0606 |
| 35 | 1620 | 22226 | 2.2226 |
| 36 | 1601 | 23827 | 2.3827 |
| 37 | 298 | 24125 | 2.4125 |
| 38 | 874 | 24999 | 2.4999 |
| 39 | 618 | 25617 | 2.5617 |
| 40 | 2640 | 28257 | 2.8257 |
| 41 | 5 | 28262 | 2.8262 |
| 42 | 149 | 28411 | 2.8411 |
| 43 | 1034 | 29445 | 2.9445 |
| 44 | 2441 | 31886 | 3.1886 |
| 45 | 460 | 32346 | 3.2346 |
| 46 | 565 | 32911 | 3.2911 |
| 47 | 1119 | 34030 | 3.403 |
| 48 | 437 | 34467 | 3.4467 |
| 49 | 927 | 35394 | 3.5394 |
| 50 | 4462 | 39856 | 3.9856 |
| 51 | 714 | 40570 | 4.057 |
| 52 | 181 | 40751 | 4.0751 |
| 53 | 1485 | 42236 | 4.2236 |
| 54 | 757 | 42993 | 4.2993 |
| 55 | 3154 | 46147 | 4.6147 |
| 56 | 2115 | 48262 | 4.8262 |
| 57 | 884 | 49146 | 4.9146 |
| 58 | 2037 | 51183 | 5.1183 |
| 59 | 1481 | 52664 | 5.2664 |
| 60 | 559 | 53223 | 5.3223 |
| 61 | 490 | 53713 | 5.3713 |
| 62 | 593 | 54306 | 5.4306 |
| 63 | 1769 | 56075 | 5.6075 |
| 64 | 85 | 56160 | 5.616 |
| 65 | 2836 | 58996 | 5.8996 |
| 66 | 213 | 59209 | 5.9209 |
| 67 | 1866 | 61075 | 6.1075 |
| 68 | 490 | 61565 | 6.1565 |
| 69 | 1487 | 63052 | 6.3052 |
| 70 | 4322 | 67374 | 6.7374 |
| 71 | 1418 | 68792 | 6.8792 |
| 72 | 1023 | 69815 | 6.9815 |
| 73 | 5490 | 75305 | 7.5305 |
| 74 | 1520 | 76825 | 7.6825 |
| 75 | 3281 | 80106 | 8.0106 |
| 76 | 2716 | 82822 | 8.2822 |
| 77 | 2175 | 84997 | 8.4997 |
| 78 | 3505 | 88502 | 8.8502 |
| 79 | 725 | 89227 | 8.9227 |
| 80 | 1963 | 91190 | 9.119 |
| 81 | 3979 | 95169 | 9.5169 |
| 82 | 1090 | 96259 | 9.6259 |
| 83 | 245 | 96504 | 9.6504 |
| 84 | 1194 | 97698 | 9.7698 |
| 85 | 994 | 98692 | 9.8692 |
| 86 | 3902 | 102594 | 10.2594 |

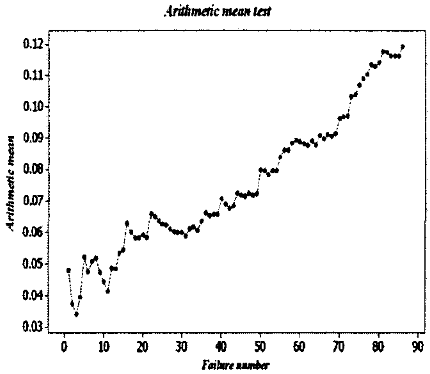

그림 2. 산술평균 검정
Figure 2. Arithmetic mean test

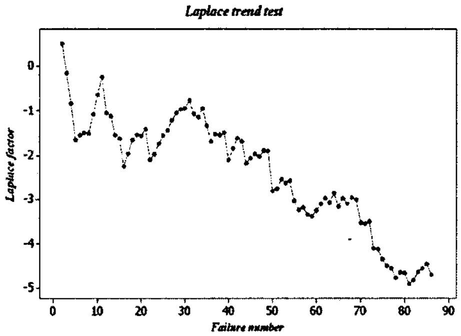

그림 3. 라플라스 추세 검정
Figure 3. Laplace trend test

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

| Model | MLE |
|---|---|
| Goel-Okumoto Model | $\hat{\beta}_{1_{MLE}}=0.5527$<br>$\hat{\theta}_{MLE}=98.3094$ |
| Yamada-Ohba-Osaki Model | $\hat{\beta}_{2_{MLE}}=0.5706$<br>$\hat{\theta}_{MLE}=89.8611$ |
| Chi-square Model(df=2) | $\hat{\sigma}^2_{MLE}=0.9371$<br>$\hat{\theta}_{MLE}=95.0458$ |
| Chi-square Model(df=4) | $\hat{\sigma}^2_{MLE}=0.9517$<br>$\hat{\theta}_{MLE}=88.1934$ |
| Chi-square Model(df=6) | $\hat{\sigma}^2_{MLE}=1.6784$<br>$\hat{\theta}_{MLE}=87.1305$ |

이 분야에서는 일반적으로 모형 선택의 하나의 방법으로 편차자승합(SSE [9])을 많이 이용하여 왔다.

본 논문은 자유도에 의한 카이제곱 모형을 비교하기 위하여 자유도와 관련된 AIC(Akaike Information Criterion) 통계량을 이용하여 비교하고자 한다. 이 통계량은 적용모형에 대하여 자유도와 관련된 우도함수를 최대화 시켜주는 척도이다. 이 척도는 다음과 같이 정의된다[9].

$$AIC = = -2\log(\max\ of\ likelihood\ \ function) + 2N$$

단, $N$ 은 적용모형에 대한 모수의 수를 의미하고 이러한 AIC 척도는 SSE와 마찬가지로 작은 값을 가지는 모형이 상대적으로 효율적인 모형으로 간주된다. 이러한 결과는 <표 3>에 요약되었다. 이 표에서 카이 제곱 분포 모형들이 자유도가 2인 경우를 제외하고는 이 분야에서 기존에 알려진 모형인 Yamada-Ohba-Osaki 모형이나 Goel-Okumoto 모형에 비해 상대적으로 효율적 모형으로 나타나고 있다.

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

| Model | SSE | AIC |
|---|---|---|
| Goel-Okumoto Model | 6057.3834 | 332.9096 |
| Yamada-Ohba-Osaki Model | 5086.1281 | 326.3824 |
| Chi-square Model(df=2) | 5409.2313 | 328.6828 |
| Chi-square Model(df=4) | 3377.8745 | 311.0964 |
| Chi-square Model(df=6) | 3167.1864 | 308.6910 |

예측 오차(prediction error)의 비정상성(nonstationarity)에 대한 측도는 Kolmogorov 거리(distance)[1,10,11]로 측정되는데 이 거리가 클수록 상대적으로 비정상성을 내포하고 있다. (그림 4)는 S-Plus 소프트웨어[12]를 이용하여 Kolmogorov 검정에 대한 그림을 보여주고 있고 이 그림에서도 전체적으로 카이제곱모형이 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 카이제곱모형의 형태가 상대적으로 비정상성 속성이 덜 내포하고 있음을 알 수 있다.

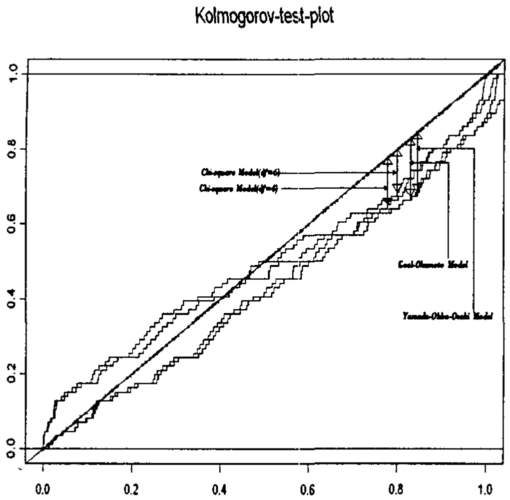


그림 4. 모형에 대한 Kolmogorov 검정 그림
Figure 4. Kolmogorov test of each model

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

| Model | Kolmogorov distance |
|---|---|
| Goel-Okumoto Model | 0.159593 |
| Yamada-Ohba-Osaki Model | 0.155121 |
| Chi-square Model(df=4) | 0.127527 |
| Chi-square Model(df=6) | 0.113803 |

## 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 발생시간 과의 관계를 효율적으로 관리함으로서 소프트웨어 신뢰도를 성장 시 킬 수 있다. 이러한 과정을 소프트웨어 성장 과정이라고 볼 수 있다.

본 논문에서는 잔존 결함 1개당 고장 발생률이 증가추세를 가진 카이제곱 분포를 이용한 카이제곱 모형을 제안하였다. 고장 간격시간으로 구성된 자료를 이용하여 기존의 모형과 카이제곱 모형에 대하여 최우 추정법을 이용하여 모수 추정을 실시하였다. 카이제곱 모형에 대한 자유도를 형상모수의 척도로 간주하여 고장수가 비교적 큰 실측 자료(고장수가 86)인 Allen P.Nikora 와 Michael R.Lyu 가 인용한 SYS2 자료을 통하여 분석하였다.

효율적인 모형 비교를 위한 편차자승합 결과 와 AIC 통계량의 결과도 카이제곱 모형이 기존에 잘 알려진 Yamada-Ohba-Osaki 모형이나[13][14] Goel-Okumoto보다 우수함을 보이고 있고 콜모고로프 거리의 결과도 카이제곱 모형이 상대적으로 정상성에 가까운 모형으로 간 주 할 수 있다. 분석된 자료에 대한 평가에서도 산술평균 검정과 라플라스 추세 검정을 실시한 결과도 신뢰성장이 되고 있음을 나타내고 있다

## 참고문헌


[1] S. S. Gokhale and K. S. Trivedi. "A time/structure based software reliability model". Annals of Software Engineering. 8, 85-121. 1999

[2] J. F. Lawless. Statistical Models and Methods for Lifetime Data. John Wiley & Sons, New York, 1981.

[3] L. Kuo and T. Y. Yang. "Bayesian Computation of Software Reliability". Journal of the American Statistical Association, Vol.91, pp.763-773, 1996.

[4] V. K, Rohatgi. Statistical inference, pages 398-416. JOHN WILEY & SONS, INCl, New York, 1984.

[5] 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.

[6] 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.
[7] A. P. Nikora and M. R. Lyu. Handbook of Software Reliability Engineering, M.R.Lyu, Editor, chapter Software Reliability Measurement Experience, pp. 255-301. MacGraw-Hill, New York, 1996.
[8] 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.
[9] 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.
[10] M. R. Lyu. Handbook of Software Reliability Engineering, M.R.Lyu, Editor, chapter Introduction, pp.3-25. MacGraw-Hill, New York, 1996.
[11] W. Q. Meeker and L. A. Escobar. Statistical Methods for Reliability Data, pages 98-101. JOHN WILEY & SONS, INCl, New York, 1998.
[12] S. Selvin. Moden Applied Biostatistical Methods Using S-Plus, pages 141-184. Oxford University Press, New York, 1998.
[13] 김희철, "일반화감마분포를 이용한 NHPP 소프트 웨어 신뢰도 모형에 관한 연구 ".한국컴퓨터정보 학회논문지, 제10권(6호), pp.27-36, 2005, 12
[14] 김홍진, "소프트웨어 재사용을 위한 소프트웨어 칩 표현식에 관한 연구".한국컴퓨터정보학회논문지, v.006, n.004, pp.12-20, 2001, 12

## 저 자 소 개

김 희 철
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